

Design and Implementation of 4K-UHD Multichannel Video Synchronization Based on MMT Protocol

Tae-Jun Jung¹, Eun-Bin Ahn², Ayoung Kim², Kwang-Eun Won², Kwang-Deok Seo²

¹Intelligent Media Lab, SK Telecom
6, Hwangsaeul-ro, Bundang-gu, Seongnam-si, Korea
tj.jung@sk.com

²Division of Computer and Telecommunications Engineering, Yonsei University
1, Yonseidaegil, Wonju-si, Gangwon, Korea
kdseo@yonsei.ac.kr

Abstract - In this paper, we propose a practical synchronization mechanism to achieve synchronized playback of the arrived 4K video bitstreams at different terminal devices based on MMT (MPEG Media Transport) protocol. In addition to the synchronized playback, we achieve the video playback with minimum service latency by estimating the accurate total end-to-end delay between the server and receiver. The conventional synchronization method between multiple clients generally calculates the time difference through communication between each client. Communication between multiple clients increases complexity and consumes network bandwidth. We do not rely on the communication between multiple clients but exploits the timestamp information supported by the MMT protocol.

Keywords: Multichannel Media Synchronization, Hybrid Video Delivery, Media Streaming, MMT Protocol.

1. Introduction

In recent years, digital broadcasting services and IP-based multimedia services over the Internet including mobile internet, have begun to be integrated and converged [1]. This trend is expected to continue with other multimedia services. In addition, emerging multimedia services and content are being introduced. These multimedia service developments mean that various services and content will be delivered over heterogeneous networks and that users will expect to consume such services using these networks, which depend on the availability and reach of the network at the time of consumption.

Along with this trend, there have been so many changes in multimedia service environments, such as media content delivery networks, diverse video signals, 4K/8K video transport systems, and various client terminals that display multi-format signals. It has become clear that the MPEG standard has been facing several technical challenges due to the emerging changes in those multimedia service environments [2]. Therefore, to address these technical challenges to existing and emerging MPEG standards, ISO MPEG developed the MPEG-H standard suite (ISO/IEC 23008) for the delivery of audio-visual information compressed with high efficiency over a heterogeneous environment.

The MPEG-H suite consists of three functional areas: HEVC (High Efficiency Video Coding), 3D audio, and MMT (MPEG Media Transport)[3]. MMT serves to address the technical challenges of existing standards due to recent changes of multimedia delivery and consumption environments and new requirements from emerging use cases and application scenarios in the area of multimedia services. MMT is intended to overcome the current limitations of available standards for media streaming by providing a streaming format that is transport- and file-format-friendly, cross-layer optimized between the video and transport layers, error-resilient for MPEG streams, and convertible between transport mechanisms and content adaptation to different networks [4]. Especially, MMT is designed to support hybrid delivery using broadcast network and the Internet for multimedia services and to support the incorporation of companion devices for additional services.

In this paper, we devise a media synchronization mechanism for hybrid video delivery of 4K-UHD multichannel video over heterogeneous environments. Media synchronization is concerned with maintaining the requirements of the temporal relationships between two or more media [5]. To support synchronization between multichannel media transported over hybrid networks, the MMT protocol is employed in this paper.

2. Hybrid Delivery of 4K-UHD Multichannel Video Using MMT Protocol

Hybrid delivery is defined as simultaneous delivery of one or more content components over more than one different types of network. Basic concept of hybrid delivery is to combine media components on different channels. One example is that one media component is delivered on broadcast channels and the other media component is delivered on broadband networks.

Figure 1 represents the structure and the interface of the transmission module used for hybrid delivery of MMT protocol-based 4K-UHD multichannel video. This module extracts an NAL stream by parsing the compressed H.264 or MP4 video files to make them suitable for real-time streaming. Subsequently, in order to transmit the extracted NAL stream to the receiving end, the stream is packetized in the MMTP payload module according to the three fragment types: MPU (Media Processing Unit) metadata, fragment metadata, and MFU (media fragment unit), and then transmitted to the MMTP packetizing module [3]. When it is delivered to the MMTP packetizing module, the information needed for efficient packetization and delivery of media data, such as the NTP timestamp, is packetized according to the Asset ID. In the case of Fig. 1, four different multichannel videos, each of which corresponds to a specific view, are generated and transmitted in hybrid delivery way using the MMT protocol.

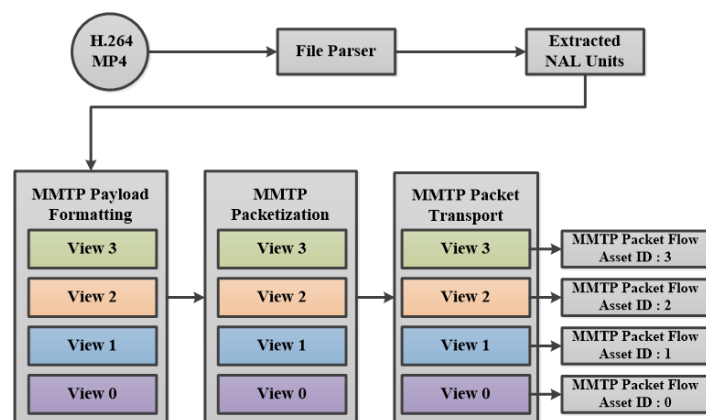


Fig. 1: 4K-UHD multichannel video sender module using MMT protocol.

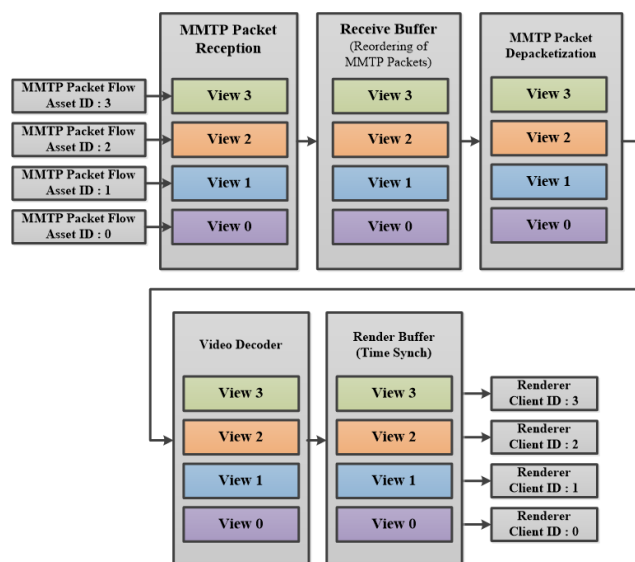


Fig. 2: 4K-UHD multichannel video receiver module using MMT protocol.

Figure 2 represents the structure of the MMT protocol-based 4K-UHD multichannel video receiving module and relationship between various interfaces. If no abnormality is detected, the packet is stored in the reception buffer. After

removing the header of MMTP packet stored in the reception buffer, the MMTP packet is then delivered to the MMTP payload depacketizing module for MMTP depacketization. In the MMTP payload depacketizing module, after stripping off the MMTP payload header, the NAL unit information is delivered to the decoder based on the metadata information. When the decoder receives the NAL unit information, the information is stored in the render buffer after decoding and the synchronization is adjusted based on the MMT timestamp information. Thereafter, the reconstructed four videos are presented on the screen in synch with each other.

3. Synchronization of the Multichannel Videos Based on MMT Protocol

For those transmitted multichannel video streams, it is necessary to provide synchronized playback at the receiver as shown in Fig. 2. In the developed MMT protocol-based hybrid delivery of 4K-UHD multichannel video, synchronization is achieved by utilizing the MMT timestamp field which is originally used to calculate network jitter instead of using the MPU timestamp descriptor information used for providing initial presentation time of each MPU. This approach could avoid a waste of bandwidth caused by the excessive transmission of the signaling message containing MPU timestamp descriptor information. Figure 3 shows the modified structure of MMT protocol packet header adopted for the implementation. Although the timestamp included in the original MMT protocol packet header is designed to indicate the delivery instance of the packet and calculate delay jitter experienced during the packet transmission, we modified the timestamp to indicate presentation time for the payload data of each MMT Asset.

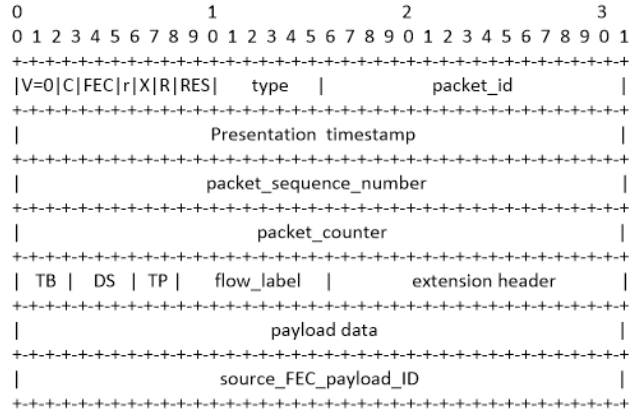


Fig. 3: Modified MMT protocol packet header adopted for the implementation.

Furthermore, in order to obtain precise synchronization, it is necessary to adjust the clock mismatch between the server and the client [6]. Figure 4 shows how to synchronize the actual playout start time between the server and the client when the initial clock mismatch between the two terminals is 10ms for example. First, client sends a connect-request message which includes a delivery timestamp (denoted by A in Fig. 4) in an MMT signaling message to the server. Then the server sends a reply message with a timestamp value (denoted by B in Fig. 4) to indicate the time instant when the server received the client's connect-request message. When the reply message containing B value is arrived at the client, the client measures the arrival time of the reply message (denoted by C in Fig. 5-8) and uses it together with A and B to calculate the clock mismatch between the server and the client. The *delta* time denoting the clock mismatch between the server and the client is calculated by the following equation:

$$\delta = C - \frac{RTT}{2} - B \quad (1)$$

This routine process iterates 10 or more times to obtain sufficient accuracy for the RTT between the server and the client. The client calculates and updates RTT time for each routine, and then calculates averaged RTT time. Using the *delta* time obtained from Eq. (1), actual latency time can be expressed by adding the *delta* time to the desired latency time. So the actual playout start time in the case of Fig. 4 is determined to be 10:04:10.

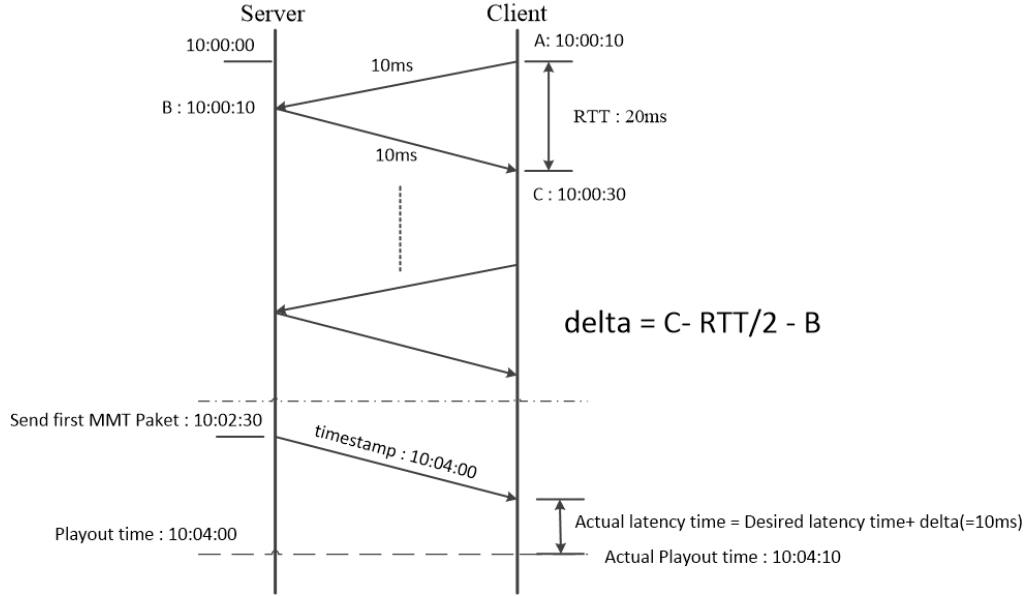


Fig. 4: How to synchronize the actual playout start time between the server and the client (clock mismatch of 10ms case).

4. Experimental Results

We implemented 4K-UHD multichannel video composition and synchronization scheme based on MMT protocol. We used C++, FFmpeg library [7], and SDL2 library. A single 8K resolution video is divided into four 4K-UHD multichannel videos (24fps, yuv 4:4:4), so that the server can send the four 4K-UHD multichannel videos to four different clients through different IP channels. Each client receives only one 4K-UHD video stream corresponding to a specific view. To each 4K-UHD video stream, different Asset ID is assigned to discriminate it from other 4K-UHD video streams. Figure 5 compares the synchronization mismatch in millisecond unit for each frame between the required playout time and the actual playout time at the four clients. According to the experimental results, the synchronization mismatch is ranged between maximum of 31 ms and minimum of 0 ms. This result corresponds to the synchronization mismatch within just one video frame for each client. Figure 6 shows the snap shot of the composed 4K-UHD multichannel videos displayed at the instant of 15 seconds since the start of the playout. From the results, it is confirmed that accurate synchronization among the four clients could be achieved by the proposed mechanism.

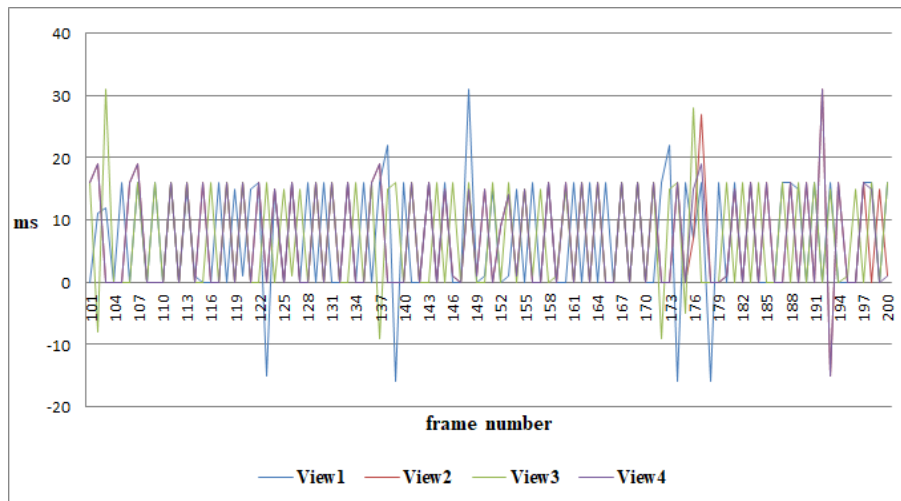


Fig. 5: Comparison of averaged synchronization mismatch for each frame at the four clients (for frames 101 through 200).



Fig. 6: Snap shot of the composed full 8K-UHD multichannel video displayed at 50 seconds since the start of the playout.

5. Conclusion

We proposed a practical synchronization mechanism to achieve synchronized playback of the arrived four 4K-UHD video bitstreams at different terminal devices. To achieve the media synchronization, we exploited the MMT technology that provides basic synchronization tools to support media synchronization over hybrid delivery environments. In order to resolve the clock mismatch between the server and the client, the actual playout time is adjusted by using the proposed δ value considering the RTT. This results in an accurate clocking operation between the different clients. According to the experimental results, the synchronization mismatch is ranged between maximum of 31 ms and minimum of 0 ms. We also achieve the video playback with minimum service latency by estimating the accurate total end-to-end delay between the server and the client.

Acknowledgements

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-2015R1D1A1A01058873).

References

- [1] S. Aoki, K. Aoki, H. Hamada, Y. Kanatsugu, M. Yamamoto, and K. Aizawa, "A new transport scheme for hybrid delivery of content over broadcast and broadband," in *Proceedings of the IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*, 2011.
- [2] Y. Lim, K. Park, J. Lee, S. Aoki, and G. Fernando, "MMT: An Emerging MPEG Standard for Multimedia Delivery over the Internet," *IEEE Multimedia*, vol. 20, no. 1, pp. 80-85, 2013.
- [3] ISO/IEC 23008-1, *High efficiency coding and media delivery in heterogeneous environments – MPEG-H Part 1: MPEG Media Transport (MMT)*, 2014.
- [4] Y. Lim, S. Aoki, I. Bouazizi, and J. Song, "New MPEG transport standard for next generation hybrid broadcasting system with IP," *IEEE Trans. Broadcasting*, vol. 60, no. 2, pp. 160-169, 2014.
- [5] S. Din and D. Bulterman, "Synchronization techniques in distributed multimedia presentation," in *Proceedings of the Fourth International Conference on Advances in Multimedia (MMEDIA)*, pp. 1–9, 2012.
- [6] T. Jung and K. Seo, "A client-driven media synchronization mechanism for RTP packet-based video streaming," *Journal of Real-time Image Processing*, vol. 12, no. 2, pp. 455-464, July 2016.
- [7] *FFmpeg library*. [Online]. Available: <https://www.ffmpeg.org/download.html>.