

일반논문 (Regular Paper)

방송공학회논문지 제24권 제2호, 2019년 3월 (JBE Vol. 24, No. 2, March 2019)

<https://doi.org/10.5909/JBE.2019.24.2.299>

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

하이브리드 미디어 전송을 위한 MMT 프로토콜 기반의 대역내 미디어 동기화 기법

원 광 은^{a)}, 안 은 빙^{a)}, 김 아 영^{a)}, 이 흥 래^{a)}, 서 광 덕^{a)}‡

In-band Media Synchronization Mechanism Based on MMT Protocol for Hybrid Media Delivery

Kwang-eun Won^{a)}, Eun-bin Ahn^{a)}, Ayoung Kim^{a)}, Hong-rae Lee^{a)}, and Kwang-deok Seo^{a)}‡

요 약

ISO MPEG에서는 향후 대용량의 초고해상도 영상을 이종(Heterogeneous)의 네트워크 환경에서도 소비자에게 원활하게 전달하기 위해 MPEG-H를 표준화 하였다. 그 중 MMT(MPEG Media Transport) 프로토콜은 이종의 네트워크 환경에서 하이브리드 (hybrid) 미디어 전송을 할 수 있는 차세대 MPEG 미디어 전송 규격이다. 대용량 초고해상도의 영상을 이종의 네트워크 환경을 통해 원활하게 서비스하려면 전송된 미디어 간의 정확한 동기화 (synchronization)가 중요한 요소이다. 본 논문에서는 MMT 기술에서 제공하는 대역 외 (out-of-band) 미디어 동기화 기법의 추가적인 세션 연결과 과도한 트래픽 발생 문제를 해결하기 위해 대역내 (in-band) 미디어 동기화 기법을 제안하고 그 성능을 검증한다.

Abstract

In ISO MPEG, MPEG-H has been standardized to smoothly deliver huge volume of ultra-high resolution videos over heterogeneous network environments for the consumers. In the MPEG-H standard, MMT(MPEG Media Transport) protocol is the next generation MPEG media transport technology to provide hybrid media delivery over heterogeneous environments. It is important to sustain accurate media synchronization to provide ultra-high resolution video services over heterogeneous network environments. In this paper, we propose and verify an in-band media synchronization scheme to circumvent the problem of additional session setup and excessive traffic problems of the conventional out-of-band media synchronization supported by the MMT technology.

Keyword : in-band media synchronization, hybrid media delivery, MMT protocol

a) 연세대학교 컴퓨터정보통신공학부(Division of Computer and Telecommunications Engineering, Yonsei University)

‡ Corresponding Author : 서광덕(Kwang-deok Seo)

E-mail: kdseo@yonsei.ac.kr

Tel: +82-33-760-2788

ORCID: <https://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-2018R1D1A1B07047065).

· Manuscript received December 3, 2018; Revised March 2, 2019; Accepted March 2, 2019.

I. 서 론

ISO MPEG에서는 UHD (ultra high definition) 해상도의 비디오를 IP 망을 포함하는 이종의 네트워크 환경에서 하이브리드 전송 방식을 통해 안정적으로 전달하기 위한 MPEG 표준 기술 개발의 필요성과 중요성을 인식하여 MMT (MPEG Media Transport) 표준화 작업을 2011년부터 진행해 오고 있고, 2015년에 1차적인 표준화 작업을 완료하였다^[1]. MMT 프로토콜 (ISO/IEC 23008-1)은 지난 20년간 멀티미디어 전송 프로토콜로서 활용되었던 RTP 및 MPEG-2 TS 시스템 기술을 대체할 수 있는 차세대 미디어 전송 기술로 간주되어 유무선 인터넷 환경에서의 실시간 미디어 전송에 활용되기 시작하였고, ATSC 3.0과 같은 차세대 디지털 방송 규격의 표준 방식으로 채택되었을 뿐만 아니라 이미 일본에서는 NHK에 의해 8K-UHD 기반의 차세대 방송을 위한 전송 방식으로 활용이 되고 있다^[2]. 특히, 최근 몇 년간 몰입형 (immersive) 미디어 서비스의 비약적 발달로 AR(Artificial Reality) 및 VR(Virtual Reality)과 같은 대용량 초고해상도 미디어를 실시간으로 전송하는 서비스에 MMT 프로토콜이 활용되고 있다^{[3][4]}.

MMT 기술은 기본적으로 그림 1과 같은 단대단 (point-to-point) 송수신 모델을 만족시키기 위해 필요한 미디어 인

캡슐레이션 포맷인 MPU (Media Processing Unit) 및 신호 전달을 위한 시그널링 메시지 포맷 그리고 MMT 패킷 전송을 위한 프로토콜 규격을 포함한다. 그림 1에서 패키지 (Package)는 서비스 되는 특정 프로그램에 포함되는 다양한 미디어들을 통합하는 논리적인 실체에 해당되며, 특정 미디어에 해당하는 하나 이상의 Asset을 패키지에 포함한다. 특정 Asset을 구성하는 미디어는 필요할 경우 부호화 과정을 거쳐서 MPU로 생성되는데, MPU는 독립적인 미디어 소비가 가능한 최소의 데이터 단위이다. MPU는 ISO BMFF (Base Media File Format) 파일포맷과 호환을 갖는다. 미디어 내에서 생성되는 패킷 데이터 간의 시간적 동기화는 ISO BMFF에서 기본적으로 제공하는 PTS (presentation timestamp)를 활용할 수 있다. 그러나, 이러한 PTS 시간 정보는 특정 미디어 스트림 내에서의 상대적으로 시간 간격에 의해 표현되므로 서로 다른 소스에서 생성되고 서로 다른 채널을 통해 전송되는 미디어들 간의 동기화에는 직접적으로 활용할 수 없다. 이러한 문제를 해결하기 위해 MMT에서는 대역외 (out-of-band) 방식의 미디어 동기화 기법을 제공한다. 그러나 이 방식은 추가적인 시그널링 메시지 전달에 필요한 MMT 세션 연결과 네트워크 대역폭의 낭비를 초래할 수 있다. 본 논문에서는 이러한 문제를 해결하기 위하여 미디어간 동기화에 필요한 정보를 MMT 프로

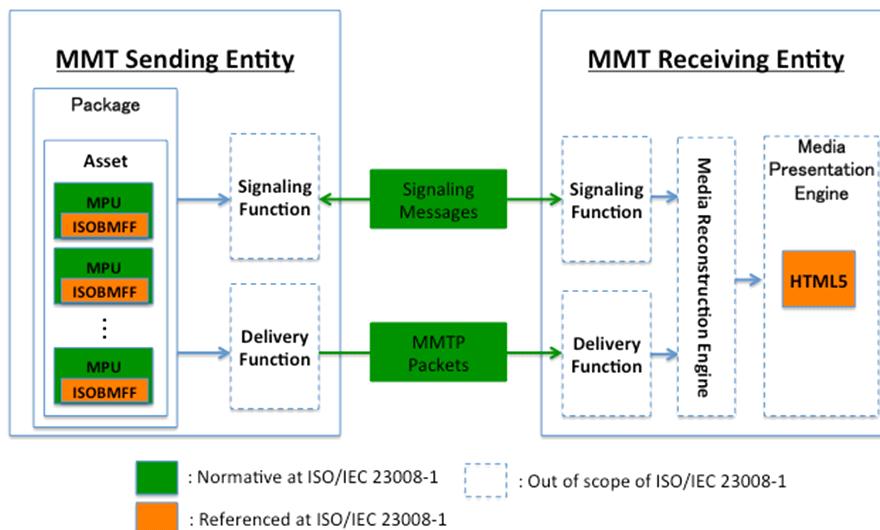


그림 1. 단대단 (point-to-point) 단일 채널 송수신 모델에서의 MMT 기술의 적용
Fig. 1. Application of MMT to point-to-point transceiver model over a single channel

토콜 헤더에 기록하여 페이로드 데이터와 함께 전송함으로써 추가적인 MMT 세션 연결과 대역폭 낭비 없이 동기화를 제공하는 대역내 미디어 동기화 방법을 제안한다.

II. MMT 기술의 대역외 미디어 동기화 기법 개요

MMT 기술에서는 서로 다른 미디어 스트림 간에 동기화 정보를 제공할 수 있는 시그널링 기법으로써 MPU_timestamp_descriptor를 제공한다^[1]. MPU_timestamp_descriptor는 특정 Asset의 MMT 미디어 스트림에 포함되는 MPU들의 초기 재생시작 시간 (Initial Presentation Time)을 UTC (Universal Time Coordinated) 시간으로 표시하여 시그널링 메시지를 통해 전송 할 수 있다. 그림 2는 MPU_timestamp_descriptor의 역할과 신택스 (syntax) 구조를 나타낸다.

그림 2에서 Asset#1과 Asset#2에 각각 포함되는 MPU#1들이 서로 동기화 대상이며, 또한 MPU#2 들이 서로 동기화 대상이다. MPU_timestamp_descriptor에 포함되는 중요 필드로는 mpu_sequence_number와 mpu_presentation_time

이 있다. mpu_sequence_number는 Asset 내에서 현재 MPU의 순서번호를 나타내고, mpu_presentation_time은 현재 MPU에 포함되어 있는 첫 번째 미디어 샘플의 재생 시작 시간을 UTC 시간으로 표시한다. 송신 측에서는 이러한 대역외 형태의 시그널링을 수신 측에 전송하고 수신 측에서는 mpu_presentation_time을 통해 MPU의 초기 재생 시간을 알 수 있고, 서로 다른 Asset에 포함되는 MPU들 간의 동기화에 활용할 수 있다. 그러나 위 방법은 대역외 시그널링 방식이기 때문에 이 신호를 목적지로 전달하기 위해서는 추가적인 MMT 세션 연결 설정과 네트워크 대역폭 낭비가 유발될 수 있다. 그러므로 네트워크에 더 이상의 부하를 주지 않도록 MMTP 패킷의 헤더 필드에 필요한 정보를 기록하여 미디어 데이터 정보와 함께 전송하는 대역내 미디어 동기화 방식의 개발이 필요하다.

III. 제안하는 MMT 프로토콜 기반의 대역내 미디어 동기화 기법

오디오 데이터는 디코딩된 오디오 데이터의 표본율 (sampling rate)을 참조하여 렌더러(renderer)에 전송된 후

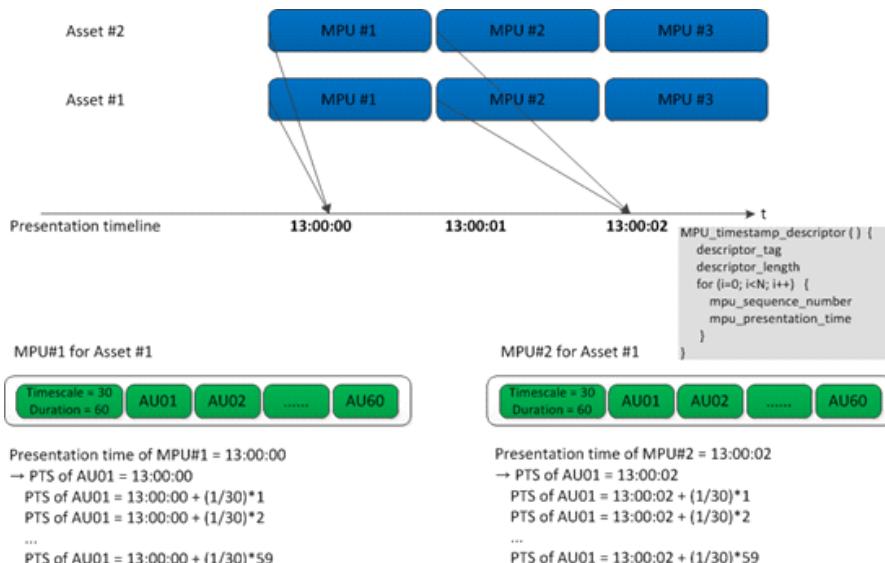


그림 2. MMT MPU_timestamp_descriptor 메시지 기반의 미디어 동기화 원리
Fig. 2. Synchronization of MMT based on MPU_timestamp_descriptor message

재생되고 비디오 데이터는 디코딩된 비디오 데이터를 화면율(frame per second)과 타임베이스 (timebase) 정보에 맞춰 비디오 렌더러에 전송된 후 재생된다. 만약 디코딩된 오디오와 비디오 데이터의 동기화 과정 없이 각각의 렌더러로 전송해서 재생하면 각 미디어 데이터의 출력률 차이 때문에 동기화가 이루어지지 않은 상태로 재생된다. 본 연구에서는 대역내 미디어 동기화를 제공하기 위하여 MMT 전송 프로토콜 헤더에 존재하는 64비트의 타임스탬프 필드를 활용한다. 이 필드는 네트워크에서 발생되는 지터(jitter)와 전송 지연 시간을 계산하는데 활용될 수 있지만, 본 연구에서는 그림 3과 같이 대역내 동기화에 필요한 PTS 시간 기록을 위해 이 필드를 활용하였다. 일반적인 ISO BMFF 파일의 moov 헤더에 기록되는 PTS 시간 정보와는 달리 미디어 재생 시간을 UTC 시간 포맷으로 표기하여 서로 다른 채널을 통해 하이브리드 방식으로 전송된 미디어들 간에 동기화가 가능하도록 설계하였다.

그림 3. 대역내 미디어 동기화를 위한 PTS 정보를 포함한 MMT 프로토콜 헤더의 구조

Fig. 3. MMT protocol header containing PTS for in-band media synchronization

PTS 정보를 기반으로 오디오 데이터 패킷과 비디오 데이터 패킷 간의 동기화를 맞추는 방법은 크게 세 가지로 나뉜다. 첫째, 비디오 데이터를 오디오 데이터에 맞춰 동기화를 맞추는 방법, 둘째, 오디오 데이터를 비디오 데이터에 맞춰 동기화를 맞추는 방법, 셋째, 외부 시스템 클럭에 오디오 데이터와 비디오 데이터의 동기화를 맞추는 방법이 있다.

본 논문에서는 첫 번째 방법을 기반으로 동기화를 맞추는데, 그 이유는 크게 두 가지이다. 첫째, 오디오의 경우 비디오와 달리 동기화 과정에서 소리의 끊김이 발생할 경우 주관적 품질에 좋지 못한 영향을 보다 더 크게 미칠 수 있지만, 비디오의 경우 동기화를 위해 화면간 시간 간격 (inter-frame interval)을 조절하더라도 이 간격의 변화를 인간의 시각이 쉽게 인지하지 못한다. 둘째, 오디오의 경우 데이터 볼륨이 비디오에 비해 매우 작기 때문에 동일한 시간 동안의 데이터에 대해 전송 및 수신 후 신호처리에 소요되는 지역 시간이 비디오 보다 짧다. 따라서, 오디오 데이터의 수신 지역으로 발생되는 문제는 거의 없고, 대용량 전송으로 늦게 수신되는 비디오 데이터의 화면 재생률 조절을 통해 오디오와의 동기화를 맞추기가 용이하다.

그림 4는 본 논문에서 제안하는 MMT 프로토콜 헤더의 UTC 포맷의 PTS 정보를 활용하는 대역내 미디어 동기화 기법의 동작 절차이다. 수신 측에서 오디오 패킷과 비디오 패킷을 수신한 후 각 패킷 헤더에 저장된 PTS 정보와 시스템 클럭(system clock)을 활용하여 두 미디어 데이터의 실제 지연 차이(actual delay)를 구하고 비디오 데이터를 오디오 데이터에 동기화 시키는 방식으로 두 미디어 스트림 간에 동기화를 수행한다. 송신 측에서 저장한 오디오 데이터 패킷과 비디오 데이터 패킷의 PTS 정보 및 타임베이스 정보를 수신 측에 있는 Delay Generator에 전달한다. Delay

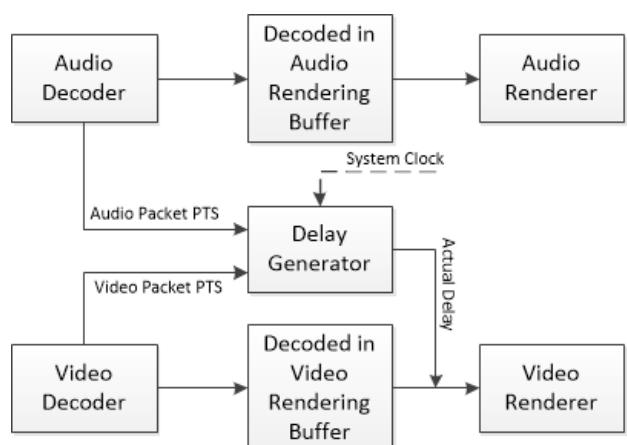


그림 4. 제안하는 오디오 및 비디오 스트림 간 동기화 동작 절차

Fig. 4. Proposed Synchronization process between audio and video streams

Generator에서는 전달 받은 정보에 현재의 시스템 클럭을 참조하여 현재 프레임으로부터 다음 프레임이 재생되기 까지의 실제 지연 시간(delay time)을 계산하고 비디오 렌더러에 이를 전달해 비디오 화면 간 재생 시간 간격을 조절함으로써 오디오와 비디오 데이터 간에 동기화를 맞춘다.

IV. 실험결과

인터넷을 통한 하이브리드 미디어 전송 환경에서 동기화 성능을 실험하기 위하여 두 대의 서버를 활용하였고, 비디오 데이터 전송 서버와 오디오 데이터 전송 서버를 서로 다른 위치에 설치하였다. 각각의 서버는 MMT 표준 규격에 따른 세션 세팅 절차를 통해 하나의 동일한 클라이언트와 동시에 MMT 세션을 맺었다. 비디오의 경우 카메라에 캡처된 실시간 4K 해상도 영상을 FFmpeg으로 압축한 후 MMT 패킷에 실어서 클라이언트로 전송하였고, 오디오의 경우 마이크에 입력된 PCM 신호를 무압축으로 MMT 패킷에 실어서 클라이언트로 전송하였다. MMT 패킷 헤더에는 각각의 미디어 탑업의 데이터율에 맞게 PTS 시간 정보가 UTC 포맷으로 기록되도록 하였다. 물리적으로 분리된 채널을 통해 MMT 패킷에 실려서 전송된 비디오 및 오디오 데이터는 III장에서 제안하는 대역내 미디어 동기화 기법에 의해 수신측에서 동기화 되었다. 이 실험에서는 비디오 카메라의 화면 캡처율을 33 fps(frames/second)로 설정하였고, 대용량 비디오 전송에 의한 네트워크에서의 전달 지연과 클라이언트 컴퓨터 성능의 한계로 인한 처리 지연으로 발생

하는 비디오 화면 멈춤 현상이 발생하는 경우 화면간 시간 간격 조절을 통해 동기화를 실시하였다. 그림 5는 이 실험에서 오디오와의 동기화를 성취하기 위해 비디오의 화면간 시간 간격을 10ms로 설정한 경우의 비디오의 화면간 시간 간격 변화 추이를 나타낸다. ①번 구간 전 까지는 정상적으로 영상이 출력되고 있으며 화면간 시간 간격이 30ms에 정상적 수치에 근접해 있는 것을 확인할 수 있다. ①번 구간은 비디오 데이터의 수신 및 처리 지연으로 2.95초 동안 화면이 멈추어 있는 구간이며 ②번 구간은 2.95초 동안 재생하지 못했던 프레임들을 10ms의 화면간 시간 간격으로 빠르게 재생하여 오디오와 동기화를 맞추는 구간이다. 이후 영상이 원래의 화면율로 재생되어 정상적인 화면간 시간 간격을 유지함을 확인할 수 있다.

그림 6은 6.5초 이후에 화면이 멈추게 되는 경우 다양한 재동기화 지연 (resynchronization delay) 값에 따른 동기화 왜곡 (synchronization skew)과 동기화 수렴 속도를 나타낸다. 재동기화 지연은 현재 영상 프레임으로부터 다음 프레임까지의 시간적 간격을 짧게 주기 위한 상수 값으로서 오디오 스트림과 비디오 스트림간의 동기화가 어긋났을 때, 재동기화를 맞추기 위하여 오디오 스트림과 동기화가 맞을 때까지 화면간 시간 간격 대신에 사용되는 값이다. 동기화 왜곡은 비디오 스트림과 오디오 스트림 간의 재생시간에 있어서의 차이를 의미한다. 화면이 일정시간 동안 멈추고 오디오만 계속해서 재생되는 상황에서 재동기화 지연 값이 클수록 재동기화를 맞추는 시점이 늦어지는 것을 확인할 수 있다. 지연 제약형 (delay constrained) 비디오 전송 서비스를 위해 요구되는 최적의 지연 값 도출을 통해 재동기화 지연 값을 선택할 수도 있다^[5].

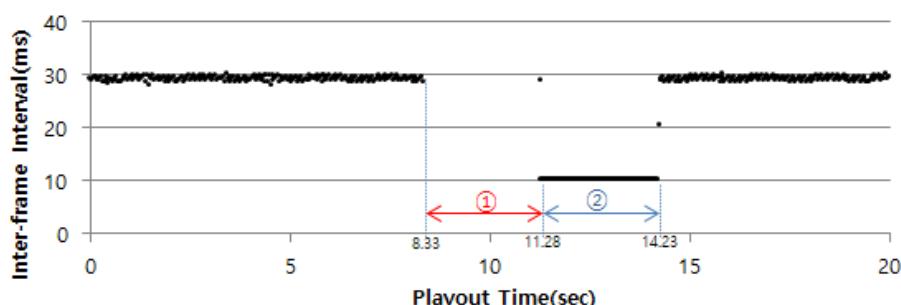


그림 5. 화면간 시간 간격 조절을 통한 동기화 성능을 나타내는 그래프.
Fig. 5. Synchronization performance achieved by adjusting inter-frame interval

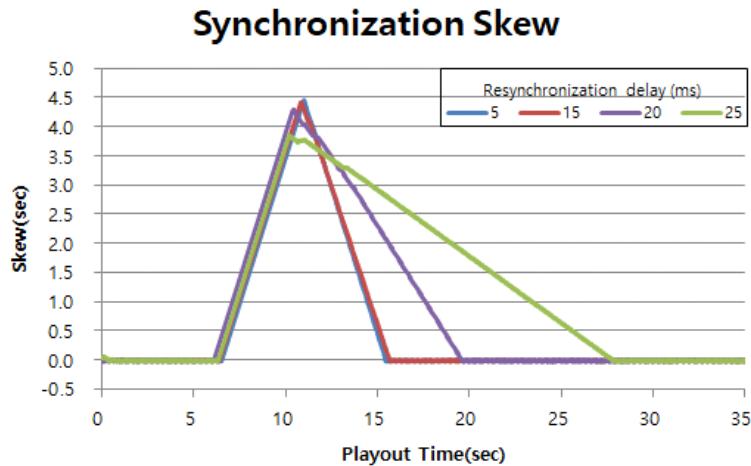


그림 6. 재동기화 지연 값에 따른 동기화 왜곡을 나타낸 그래프
Fig. 6. Synchronization skew affected by the resynchronization delay

V. 결 론

본 논문에서는 MMT 기술에서 기본적으로 제공하는 대역외 (out-of-band) 미디어 동기화 기법의 추가적인 세션 연결과 과도한 트래픽 발생 문제를 해결하기 위한 대역내 (in-band) 미디어 동기화 기법을 제안하였다. 제안 기법에서는 MMTP 패킷 헤더에 UTC 포맷의 PTS 정보를 기록하여 패킷을 전송하게 되고, 오디오 스트림을 기준으로 비디오 스트림의 화면간 시간 간격을 조절함으로써 동기화를 맞출 수 있음을 실제 구현을 통해 확인하였다. 또한, 재동기화 지연 값을 쉽게 설정함으로써 동기화 수렴속도를 높일 수 있었다.

참 고 문 헌 (References)

- [1] ISO/IEC 23008-1:2014 (First edition), Information technology High efficiency coding and media delivery in heterogeneous environments Part 1: MPEG media transport (MMT), Jun. 2014.
- [2] Y. Sohn, M. Cho, and J. Paik, "Design of MMT-based broadcasting system for UHD video streaming over heterogenous networks," Journal of Broadcast Engineering, Vol. 20, No 1, pp. 16-25, Jan. 2015.
- [3] Y. Hu, S. Xie, and Y. Xu, "Dynamic VR live streaming over MMT," in Proc. of IEEE Int. Symp. Broadband Multimedia Systems and Broadcasting (BMSB), pp. 11-16, 2017.
- [4] S. Din and D., Bulterman, "Synchronization techniques in distributed multimedia presentation," in Proc. of Int. Conf. on Advances in Multimedia, pp. 1 - 9, 2012.
- [5] C. Kim, H. Lee, T. Jung, B. Kim, and K. Seo, "An efficient delay-constrained ARQ scheme for MMT packet-based real-time video streaming over IP networks," Journal of Real-time Image Processing, Vol. 12, No. 2, pp. 257-271, Aug. 2016.

저자 소개



원광은

- 2016년 2월 : 연세대학교 컴퓨터정보통신공학부 학사
- 2018년 8월 : 연세대학교 전산학과 석사
- ORCID : <https://orcid.org/0000-0002-6370-0608>
- 주관심분야 : 영상통신, 멀티미디어 통신시스템



안은빈

- 2016년 8월 : 연세대학교 컴퓨터정보통신공학부 학사
- 2017년 3월 ~ 현재 : 연세대학교 전산학과 석박통합과정
- ORCID : <https://orcid.org/0000-0001-7681-4682>
- 주관심분야 : 영상처리, 영상통신, 360/VR 영상



김아영

- 2016년 2월 : 연세대학교 컴퓨터정보통신공학부 학사
- 2016년 3월 ~ 현재 : 연세대학교 전산학과 석박통합과정
- ORCID : <https://orcid.org/0000-0002-3793-1365>
- 주관심분야 : 영상통신, 실시간 스트리밍, 360/VR 영상



이홍래

- 2010년 2월 : 연세대학교 컴퓨터정보통신공학부 학사
- 2012년 8월 : 연세대학교 전산학과 석사
- 2012년 9월 ~ 현재 : 연세대학교 전산학과 박사과정
- ORCID : <https://orcid.org/0000-0001-6985-6016>
- 주관심분야 : 영상부호화, 영상통신, 멀티미디어 통신 프로토콜



서광덕

- 1996년 2월 : KAIST 전기 및 전자공학과 학사
- 1998년 2월 : KAIST 전기 및 전자공학과 석사
- 2002년 8월 : KAIST 전기 및 전자공학과 박사
- 2002년 8월 ~ 2005년 2월 : LG전자 단말연구소 선임연구원
- 2012년 9월~2013년 8월 : Courtesy Professor, Univ. of Florida, USA
- 2005년 3월 ~ 현재 : 연세대학교 컴퓨터정보통신공학부 교수
- ORCID : <http://orcid.org/0000-0001-5823-2857>
- 주관심분야 : 영상부호화, 영상통신, 디지털 방송, 멀티미디어 통신시스템