

# UWV 콘텐츠의 전송과 동기화를 위한 MMT 활용사례 연구

\*유성열 \*안효민 \*김규현

\*경희대학교

\*Sungryeul@khu.ac.kr \*hyominan@khu.ac.kr \*kyuheonkim@khu.ac.kr

## A case study on using MMT for transmission and synchronization of UWV content service

\*Rhyu, Sung-yeul \*An, Hyo-Min \*Kim, Kyu-Heon

\*Kyung Hee University

### 요약

본 논문은 UWV 서비스의 다양한 시나리오들을 소개하고 UWV서비스의 제공측 구성과 소비측 재생에 따라 개시되는 멀티미디어 콘텐츠 파일 전송의 형태를 검토한다. UWV 서비스의 전송 개시에 따라 적어도 하나 이상의 서비스 구성요소(멀티미디어 콘텐츠 파일)가 전송될 때 수신측에서의 엄밀한 동기화가 수행되지 못하면 전체 장면의 일부를 표시하는 View간의 일치가 이루어지지 않으며 이로 인해 사용자의 경험 측면에서 임장감이 낮아지거나 또는 불일치의 정도가 심한 경우 전체 장면을 이해하는데 어려움이 발생한다. 이에 본 논문에서는 기존의 전송 방식인 TS와 최근 MPEG에서 완료된 MMT를 사용해 UWV 서비스를 구성하는 콘텐츠의 전송시, 소비측에서의 동기화된 View 콘텐츠 재생을 달성하기 위한 전송 방법을 알아본다. 또한 시스템 구성을 통해 실제 전송시 요구사항의 달성여부와 함께 추가로 고려해야 할 사항들을 알아본다.

### 1. 서론

UWV(Ultra Wide Vision)란 하나 이상의 영상 취득 장치를 사용하여 Panoramic 또는 Immersive한 영상을 동시에 또는 시차를 두고 취득하여 전송과 재생을 제공하는 서비스 및 이를 구성하는 멀티미디어 콘텐츠 구성 요소를 의미한다.

UWV 서비스를 통해 소비자에게 제공하려는 영상은 자연 풍경, 경기장과 같이 선형 또는 구조적 시점을 가지고 넓게 펼쳐지는 초광역 시야(Ultra Wide Vision)이며, 하나 이상의 디스플레이 장치에서의 재생을 통해 소비자로서 하여금 실제로 그 곳에서 경험하는 것과 같은 시야를 제공하여 높은 임장감(Immersiveness)을 제공하는데 그 특성이 있다.

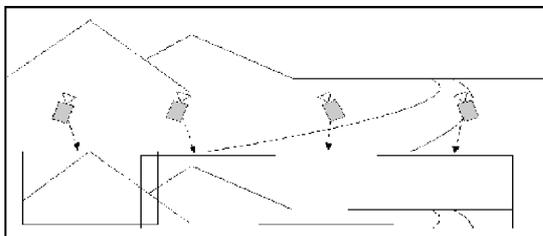


그림 1. 초광역 시야 서비스를 위한 영상 취득

UWV 서비스는 전송한 하나 이상의 영상 취득 장치를 사용하여 취득된 복수개의 멀티미디어 콘텐츠 구성 요소가 서비스 전송 개시 시점에 수신측에 전송될 때 하나의 영상 재생 장치내에서 동기화 되어 재생되거나, 복수개의 영상 재생 장치에서 동기화 되어 재생된다.

복수개 멀티미디어 콘텐츠가 동기화되어 재생되려면, 멀티미디어 콘텐츠 내의 시간 정보가 다른 멀티미디어 콘텐츠 내의 시간 정보와

연관성이 있어야 하며, 만약 연관성이 없는 경우에는 소비측에서 구성하는 하나 또는 하나 이상의 영상 재생 장치에서의 동기화 된 재생을 담보할 수 없다.

본 논문에서는 서비스 구성에 따른 UWV 멀티미디어 콘텐츠 구성요소의 전송 방법과 소비측에서의 소비 방법 그리고 이 때 동기화를 담보하기 위해 고려해야 하는 사항들을 알아보고 이를 실제 시스템을 구성해 검증한다.

### 2. UWV 서비스 시나리오

본 논문에서 UWV 서비스 제공 방법은 크게 서비스 제공 주체에 의한 집중형(Centralized)과 사용자에 의한 참여형(Participated)으로 구분한다. 집중형은 복수개 카메라에서 취득된 멀티미디어 콘텐츠의 편집과 인코딩을 UWV 서비스 서버에서 수행한다. 참여형은 복수개 카메라에서 취득된 멀티미디어가 Facebook Live, Periscope와 같이 개별 라이브 스트리밍되고 다만 UWV 서비스와의 시공간 관계(전체 UWV 영상내의 위치 등)는 UWV 서비스 서버에 의해 수행된다. 참여형의 경우 UWV 서비스에서 표현하려는 이벤트는 하나이지만, 사용자와 시점에 따라 다른 뷰가 생성되어 전송된다.

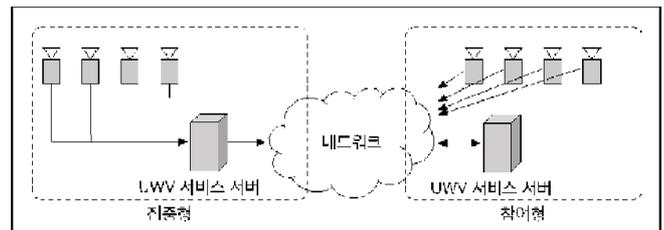


그림 2. UWV 서비스 제공 형태

또한 본 논문에서 제시하는 UWV 서비스 소비 방법으로는 하나 이상의 멀티미디어 콘텐츠로부터 전체 View를 구성하고 이 가운데 일부의 View를 디스플레이에서 표시하는 통합형(merged) 방식과 각각의 멀티미디어 콘텐츠를 디스플레이에서 표시하는 개별형(separated) 방식으로 구분한다. 통합형의 경우 가상의 UWV 통합 View가 구성되고 디스플레이에서 UWV 통합 View의 일부를 보이거나 스크롤하며 재생하게 되어 멀티미디어 콘텐츠가 제공하는 View와 디스플레이에서 표시하는 View가 일치하지 않는다. 개별형의 경우 멀티미디어 콘텐츠가 제공하는 View와 디스플레이에서 표시하는 View가 일치한다.

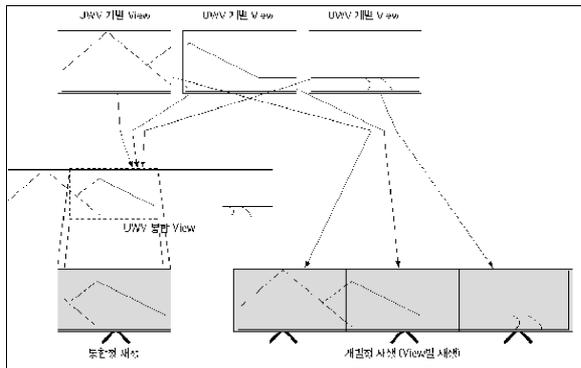


그림 3. UWV 서비스 소비 방법

### 3. UWV 서비스 전송 방법

UWV 서비스가 개시되면 서비스 제공 방법, 소비 방법에 따라 UWV 서비스의 구성요소인 멀티미디어 콘텐츠가 준비되어 전송되는데, 해당 멀티미디어 콘텐츠는 각각의 영상 취득 장치에서 사용하는 클럭을 기반으로 Timestamp를 부여받아 인코딩 된다. 즉, 영상 취득 장치간의 시간 동기화 없이 취득되고, 서비스 제공 방법에 따라 UWV 서비스 서버에서 멀티미디어 콘텐츠들을 취합 후 단일 클럭을 부여하여 인코딩 할 수도, 또는 취득 장치가 제 3의 인터넷 서비스에 개별 전송하고 UWV 서비스 서버에서는 View간 시공간 관계만을 분석하여 제공할 수도 있다.

본 논문에서는 UWV 서비스를 집중형과 참여형으로 구성하여 통합형과 개별형으로 소비할 때 TS와 MMT를 사용해 전송하는 경우, 멀티미디어 콘텐츠 재생시 동기화를 달성하기 위한 고려 사항을 검토한다.

#### 가) MPEG-2 TS

MPEG-2 TS[1]에서는 인코딩시에 인코더의 STC를 PTS로 기록하며, 개별 취득 장치의 STC가 기록된 경우와 UWV 서비스 서버의 STC가 기록된 경우로 나눌 수 있다. 또한 소비자에서 재생하기 위해 요청한 멀티미디어 콘텐츠는 각각 다른 경로를 거쳐 서로 다른 서버로부터 내려받아진다.

MPEG-2 TS의 경우, 해당 스트림이 생성되는 STC를 기반으로 독립적으로 PTS를 생성하고 있기에, 서로 다른 서버 또는 획득 장치에서 획득한 영상간의 직접적인 동기화는 불가능하다. 이에 따라, 개별적으로 획득한 영상을 모두 모아 하나의 방송 스트림으로 Multiplex하여 전송하고 수신측에서는 수신된 개별 멀티미디어 스트림간의 PTS 차이를 재생 내내 유지시키는 방법[4] 또는 하나의 STC를 통해 개별 멀티미디어 스트림의 PTS를 새롭게 부여하는 방법[5]을 사용해 서로

다른 네트워크 경로를 통해 전송된 멀티미디어 콘텐츠가 수신측에서 재생된다.

전송이 개시되면 전송 경로에 따라 멀티미디어 콘텐츠 별로 실제 전송 도착시간이 달라지게 된다. 이 때, 전체 멀티미디어 콘텐츠 중 화면 구성에 필요하여 전송된 n개의 멀티미디어 콘텐츠들의 시간과 실제 이벤트간에는 초기 지연  $Delay_{0,1,2,\dots,n}$ 이 존재한다. 하지만 MPEG-2 TS에는 실제 이벤트에 대한 시간 정보가 없으므로 수신측 기준으로 가장 처음 수신된 멀티미디어 콘텐츠간 상대지연만을 기준으로 운용하며, 이때 초기 상대지연은 멀티미디어 콘텐츠 중 화면 구성을 위해 필요한 제 1 세그먼트의 수신 완료 간 상대지연  $T\_delay_0[0], T\_delay_1[0], T\_delay_2[0], \dots, T\_delay_n[0]$ 로 나타낸다. UWV 서비스 구성을 위한 상대 지연간의 차이를 줄이기 위해서는 가장 늦게 도착한 멀티미디어 콘텐츠의 시간을 중심으로 동기화하여야 하며, 이에 따라 제 1 세그먼트 수신 완료후 지연시간  $TimeDelay(1)$ 은 다음과 같다.

$$meDelay(1) = \max(T_{delay_1}[1], T_{delay_2}[1], \dots, T_{delay_n}[1]) \dots (1)$$

$TimeDelay(1)$ 는 필요한 모든 멀티미디어 콘텐츠를 수신하기 위해 발생한 첫번째 시간 지연을 의미한다.

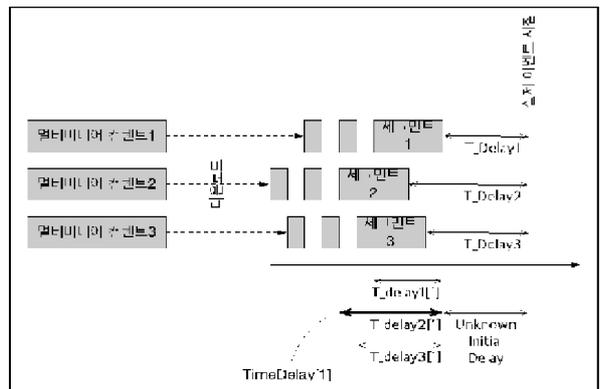


그림 4. 콘텐츠(세그먼트)간 수신 지연시간 차이

실제 이벤트 대비 지연 시간은 다음과 같다.

$$TimeDelayReal(1) = Unknown\ Initial\ Delay + TimeDelay(1) \dots (2)$$

지속되는 세그먼트 수신 과정에 다시 일부 세그먼트의 수신 지연이 발생되면 두 번째 수신완료에서 발생하는 지연시간은 다음과 같다.

$$TimeDelay(2) = \max(T_{delay_1}[2], T_{delay_2}[2], \dots, T_{delay_n}[2]) \dots (3)$$

$$TimeDelayReal(2) = UID + TimeDelay(1) + TimeDelay(2) \dots (4)$$

전송한 바와 같이 실제 이벤트와 현 PTS간의 관계가 제공되지 않으므로 실제 이벤트 대비 지연시간은 전체 세그먼트의 수신 지연이 발생할 때마다 지연 최대값만큼 지속 증가하게 된다.

$$TimeDelayReal(m) = UID + \sum_{k=0}^m TimeDelay(k) \dots (5)$$

MPEG-2 TS에는 실제 이벤트가 언제 발생하는지에 대한 정보가 없어 재생시 지연이 발생할 때마다 실제 이벤트 대비 지연시간이 누적되어 증가하는 문제가 있다. 더욱이, 한 개의 수신측 디스플레이 기기에서 UWB 서비스를 재생하는 경우에도 이와 같이 동기화 문제가 발생하기 때문에 복수개의 디스플레이 기기와 실제 이벤트 사이, 그리고 복수개의 디스플레이 기기간에 발생하는 동기화 문제를 해결하기 위해서는 별도의 방법이 필요하다.

나) MPEG-H MMT

한 편, MPEG에서는 MPEG-2 TS의 이러한 문제를 해결하고 초저지연 전송 및 동기화를 위한 전송 시스템인 MPEG-H MMT[2]가 규격화되었다.

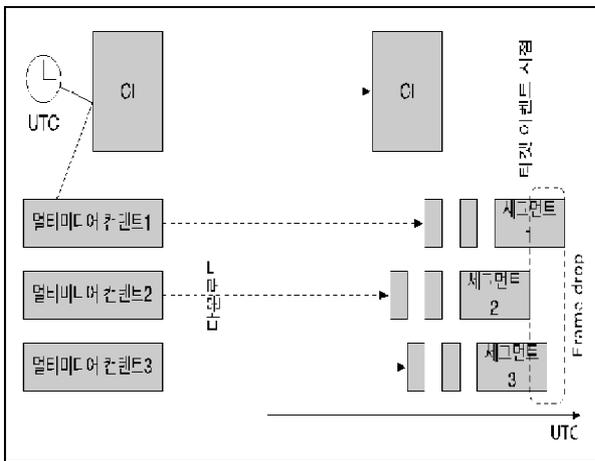


그림 5. MMT에서 UTC와 CI에 의한 타겟 시점 동기화

MMT는 멀티미디어 콘텐츠의 구성 및 전송시 미디어의 재생시간을 UTC와 동기화할 수 있는 방법을 제공한다. MMT의 전송 대상인 멀티미디어 콘텐츠의 PTS를 UTC기준의 서비스 시간에 매핑하기 위한 정보를 CI[3]라는 Manifest 를 통해 제공하고 있고, 서비스 시간이 전체 멀티미디어 콘텐츠들을 동기화하여 재생하기 위한 Target 시간이기 때문에, 소비측의 하나, 또는 여러 대의 디스플레이가 각각 UTC에 동기화되고, 전송 받은 멀티미디어 콘텐츠의 PTS 시간이 UTC로 환산되었을 때, 기기의 UTC보다 늦다면 이를 버리고 가장 근접한 UTC에 매핑되는 PTS를 가진 멀티미디어 세그먼트를 재생하도록 설계되어있다.

이에 따라, 복수개의 디스플레이 기기에서 View별 세그먼트의 동기화는 View간 동기화가 아니라 각 View와 UTC간의 동기화가 되며, 통상 NTP를 사용해 동기화 이후 기기들 간의 동기 차이는 NTP 서버와의 RTT에 따라 서로 달라진다.

기기는 주기적으로 UTC와의 동기화를 수행하기 때문에, UTC와의 동기 차이는 지속적으로 관리되며, 일부 디스플레이 기기에서 지연이 발생되더라도 Frame drop을 통해 관리되어 실제 이벤트와의 동기 차이는 누적되지 않는다.

$$meDelay_{ient1} = UTC - Clock_{Client1} \dots (6)$$

$$TimeDelay_{Client2} = UTC - Clock_{Client2} \dots (7)$$

위와 같이 하나의 UWB 서비스를 수신하는 복수개의 UWB 클라

이언트와 실제 이벤트 간의 차이는 다른 기기에 독립적으로 각 기기와 UTC와의 동기화 정확도에 기반한다. 따라서, 미디어의 재생시에 미디어 PTS를 UTC와 비교하여 계산하고, 동기화를 위해 각 Frame의 재생 시점을 결정하는 MMT 시스템의 경우, 기기간 동기화를 위한 별도의 방법이 없이도 오차의 누적 없이 실제 이벤트와의 동기 수준을 유지할 수 있다.

4. 실험 및 검증

본 논문에서는 UWB 서비스 시나리오에서 MMT를 적용한 전송 환경 구성을 통해 실험을 실시하였다.

MMT를 사용한 복수개 UWB 서버와 복수개 UWB 디스플레이 기기간 재생 및 동기화를 검증하기 위한 시스템은 다음과 같이 구성하였다.

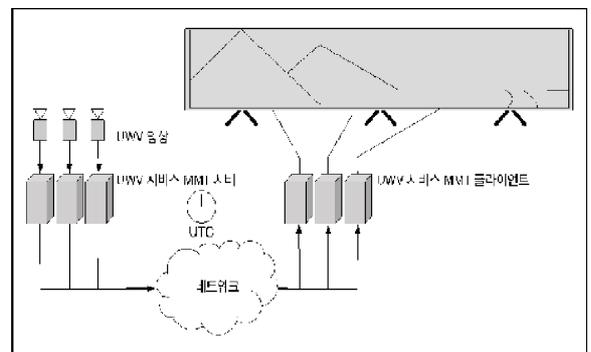


그림 6. UWB 시스템 구성예

실험은 UHD 해상도의 UWB 콘텐츠(각 3개의 UWB View)를 사용해 반복 재생시 기기별 오차가 누적되는지 여부 및 오차를 인지할 수 있는지 파악하기 위한 용도로 설계되었다.

UWB 서비스는 3개의 MMT 서버를 사용해 각각 UWB View를 별도 전송하였으며, 4K 60fps의 영상 3개가 무한 반복하여 하나의 네트워크를 통해 수신측에 전달된다.

UWB 수신측 기기들은 UTC에 동기화되며 미리 정해진 역할에 따라 3개의 UWB View중 하나를 수신하여 재생한다.

전송이 개시되면 서버로부터 CI와 영상이 주기적으로 전송되며, 수신측 기기가 수신을 시작하면, 먼저 CI를 수신하여, 전송되는 데이터로부터 멀티미디어 콘텐츠의 정보 및 수신을 시작한다. 수신측 기기는 UTC에 동기화되어있으므로, 수신 시작으로부터 수신된 멀티미디어 콘텐츠의 세그먼트의 목표 재생시간이 현재의 UTC와 얼마나 차이 나는지 확인하고 재생 또는 Frame drop을 결정한다.

한편 개별 수신측 기기에서의 재생은 다른 기기와 독립적으로 UTC에 기반하여 동기화되며 동기화의 엄밀도는 기기와 NTP 서버간의 hop과 RTT에 의존하게 된다. 일반적으로 NTP 서버와의 hop이 멀수록 오차가 커지므로 기기간의 재생시 동기화 차이가 크면 로컬 네트워크 내에 NTP 서버를 운용하는 것이 UTC 동기화 오차를 개선함에 있어 바람직하다.

또한 개별 수신측 기기의 재생이 다른 기기와 독립적이기 때문에 일 수신측 기기의 재생 시작이 전체 수신 기기들의 재생 시작과 동기화될 필요가 없다. 즉, Center view로부터 시작해 재생을 감상하다가 Left와 Right view를 나중에 시작하더라도 최초의 frame부터 전체

UWV View가 동기화되어 재생된다.

실험 결과 수일간 지속된 재생에서 동기화 오차는 누적되지 않았고, 영상 재생을 녹화해 검토한 결과 다른 2 View에 비해 최대 1 frame (60fps중 한 Frame의 차이는 16ms 오차를 의미)이 늦은 1 View가 발생하는 수준의 높은 동기화 수준이 관측되었다.

이는 재생하기 위해 수신받은 버퍼의 양과 상관없이, UTC에 시스템을 동기화하고, UTC에 기반해 현재 재생되어야 할 시점을 재생하기 때문인데, 같은 시나리오에서 MPEG-2 TS를 사용시 View를 구성하는 멀티미디어 콘텐츠간 동기화 문제가 누적되고, 기기간 동기화를 담보할 방법이 별도로 고려되어야 하는 것데 반해 MPEG-H MMT를 사용시 별도의 동기화 방법이 없이도 1 Frame 이하의 동기화 수준으로 UWV 서비스가 제공될 수 있음을 확인할 수 있었다.

## 5. 결론

이상과 같이 UWV 서비스 시나리오와 MMT를 이용한 UWV 서비스 시나리오 구현 사례에 대해 알아보았다. UWV에는 다양한 전송 시나리오가 있을 수 있는데, TS를 사용해 전송하는 경우 복수개의 클라이언트가 수신하는 환경에서 각 클라이언트간의 엄밀한 동기화 및 실제의 이벤트와 클라이언트간의 동기화등 두가지의 동기화가 요구된다. 한편 MMT를 전송시스템으로 사용하는 경우, 클라이언트와 실제의 이벤트가 모두 UTC 기반으로 동기화되기 때문에, 다른 클라이언트의 동기화 수준과 상관없이 각 클라이언트가 UTC와 최대한 동기화되는 것만으로 수 ms내의 클라이언트 및 실제 이벤트와의 동기화를 간단히 달성할 수 있었다.

향후 실험환경을 더 고도화하여 hop의 개수에 따라 복수개 클라이언트간에 발생하는 NTP 정밀도의 편차를 분석하고 로컬 네트워크 내에 로컬 NTP 서버를 세워 이에 동기화되는 클라이언트간의 NTP 정밀도에 미치는 영향을 분석하여 NTP hop과 클라이언트간 동기화 수준간의 인지 가능한 상관관계를 검토하는 것이 UWV 서비스 품질 향상에 도움이 될 것으로 생각한다.

## 참 고 문 헌 (References)

- [1] ISO/IEC 13818-1, Information Technology - Generic Coding of Moving Pictures and Associated Audio Information : Part 1 Systems, Int'l Organization for Standardization, 2007.
- [2] ISO/IEC 23008-1, "Information technology - High efficiency coding and media delivery in heterogeneous environments - Part 1: MPEG media transport (MMT)", June 2014.
- [3] ISO/IEC JTC 1/SC 29 23008-11:2014, Informaiton technology - High efficiency coding and media delivery in heterogeneous environments - Part 11 : Composition Information
- [4] J. Lee, K. Kim, H. J. Yim, W. S. Cheong, "Design and Implementation of the Higher-quality Terrestrial 3DTV Broadcasting Standard Specification Based on Synchronization with Non-Real-Time Contents," The Journal of The Korean Institute of Communication Sciences, 37(12), pp.1185-1194, December 2012
- [5] K. Yun, W. S. Cheong, G. Lee, X. LI, K. Kim, "Design of Synchronization and T-STD model for 3DTV service over Hybrid networks," ETRI Journal, volume 38, number 5, pp. 838-846, June 2016