

## 6DoF 영상 전송 구현을 위한 MPD 구조 제안

서봉석 현창종 김동호

서울과학기술대학교

(sbs91, dksshddl, dongho.kim)@seoultech.ac.kr

## MPD Structure suggestion for 6DoF video delivery

Seo, BongSeok Hyun, ChangJong Kim, Dongho

Seoul National University of Science and Technology

## 요약

최근 다양한 VR/AR HMD 장비의 등장으로 영상표준에서는 3DoF 영상인 360 영상에 관한 표준이 정립되었다. 또한 이후 추가적인 표준인 3DoF+, 6DoF 로의 표준화가 계획 중이며 진행 중이다. 하지만 현재 실질적으로 6DoF 영상을 제작하거나 전송할 수 있는 표준은 존재하지 않으며, 표준화가 진행 되더라도 전송에 적합한 형태를 찾는 과정까지 오래 걸릴 것으로 예상된다. 따라서 본 논문에서는 기존에 표준화된 360 영상 표준과 전송 표준인 DASH를 활용하여 3DoF+ 이상의 6DoF를 구현하기 위한 DASH의 MPD 구조를 제안하여 실감 미디어 전송을 할 수 있도록 한다.

## 1. 서론

최근 VR/AR HMD(Virtual Reality/Argument Reality Head Mount Display)가 시장에 공개되면서[1] 3DoF(Degree of Freedom)을 체험할 수 있는 콘텐츠에 대한 요구가 나타났다. 그리고 독립형 HMD 혹은 Hololens[2]와 같은 AR HMD의 발전으로 3DoF를 벗어나 3DoF+, 6DoF 역시 제한적이지만 체험이 가능해져, 적합한 콘텐츠에 대한 요구 역시 나타나기 시작했다. 이에 현재 VR/AR 시장은 개발 및 구현 자유도가 높은 게임을 중심으로 콘텐츠가 개발되고 있으며, 영상의 경우 360영상을 초기 콘텐츠로써 제안하고 이후 더 많은 자유도를 위한 영상 포맷, 압축방법 등에 대한 표준화를 진행하고 있다.[3] 하지만 새로운 콘텐츠의 등장에 있어서 또 하나 해결해야 할 과제는 바로 전송 표준이다. 콘텐츠의 자료 구조가 변경되어 현재 시간 단위로 잘라 전송하는 전송 표준에 적합하지 않을 수 있으며, 그에 따라 새로운 표준 콘텐츠를 전송하기 위한 방법을 표준화 하는 데에도 시간이 걸릴 것이다. 현재 다중시점 영상은 MVC (Multiview Video Coding)[4]를 통하여 압축을 실행하며, 이는 2D 영상을 기준으로 구현하고 동작할 수 있다. 따라서 현재 360 영상 구조와 전송 표준인 DASH (Dynamic Adaptive Streaming over HTTP)의 MPD(Media Presentation Description)[5]를 확장하여 3DoF 이상의 자유도를 가지는 콘텐츠가 온전히 등장하기 전 까지 공백을 메우는 기술로써 새로운 구조가 필요하다.

따라서 본 논문에서는 현재 다중시점 영상을 전달하기 위해 표준화 되었던 MVC와 영상전달 표준인 DASH에 대해 간략하게 알아본 뒤 다중시점 영상을 전달하기 위한 새로운 방법을 제시하고 그에 따른 효과를 알아볼 것이다.

## 2. 다중시점 360 영상 제공을 위한 MPD 구조 제안

## 2-1. 기존 다중 시점 영상 전송

현재 영상 전송을 위한 DASH는 그림 1 과 같이 미디어를 콘텐츠 별, 해상도별, 시간별로 분할하여 전송하는 방식을 채택하고 있다. 영상을 가장 작은 시간단위인 Segment 단위로 분할하고 이를 찾기 쉽게 Tree 형태로 구조화 한 것이다. 이때 초기에 가장 큰 시간단위를 나누는 Period가 있으며, 하위에 Adaptation Set은 콘텐츠의 종류를 나누는데 쓰이는데 소리, 자막, 영상 등을 구분할 수 있게 한다. 이때 다중시점 영상을 전송하기 위해서 기존 MVC 에서는 모든 시점을 하나의 파일로 압축하여 압축효율을 늘려 모든 시점을 한 번에 전송하는 방식을 사용하여 한 개의 Adaptation Set에 모든 시점의 영상이 포함된다. 이를 이용하여 영상을 전송하면 모든 시점의 영상을 하나의 스트림으로 전송하게 되지만 실제 사용자가 동시에 소비 가능한 시점이 한계가 있다.

## 2-2. 시점 거리에 따른 우선순위 배정

모든 시점을 하나로 압축하는 것은 압축효율 측면에서 매우 우수할 수 있으나 실제로 전송되는 전송 측면에서는 모든 시점을 전부 보내야하기 때문에 시점을 분리하여 전송하는 것이 전송대역 효율 측면에서 매우 효과적으로 작용할 것이다. 또한 사용자가 6DoF 환경에서 시점을 변경하는 것을 고려하였을 때 선택된 시점 주변부는 매우 높은 확률로 전송 받게 되지만 일정 거리 이상의 시점은 낮은 확률로 선택될 것으로 예상되기 때문에 선택된 시점을 기준으로 거리에 따라 대역 할당 비중이 달라야 한다. 그림 2와 같이 시점이 배치되어 있을 때 사

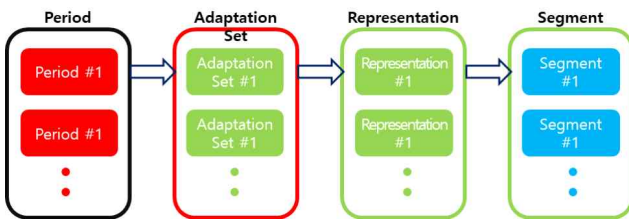


그림 1. MPD의 구조

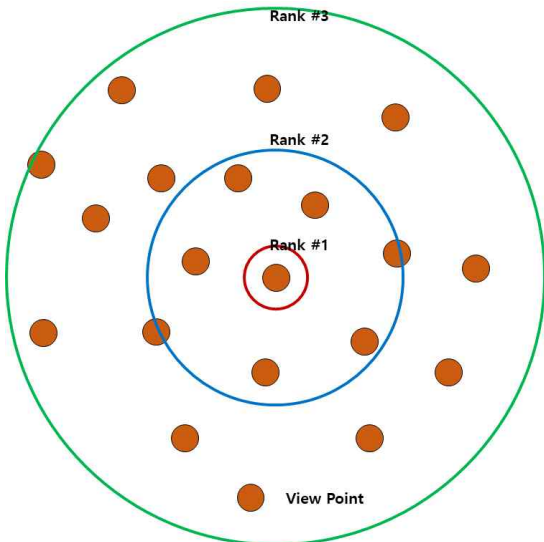


그림 2. 시점 위치에 따라 Rank가 달리 적용된 예

```
<Period>
<AdaptationSet>
<SupplementalProperty_SchemeIdUri="urn:mpeg:dash:6dof:2019" value="Rank #1">
<Role schemeIdUri="urn:mpeg:dash:role:2011" value="main"
<Representation id="1" width="3840" height="2160" ....>
<BaseURL>A.mp4
</BaseURL>
</Representation>
</AdaptationSet>
<AdaptationSet>
<SupplementalProperty_SchemeIdUri="urn:mpeg:dash:6dof:2019" value="Rank #2">
<Role schemeIdUri="urn:mpeg:dash:role:2011" value="supplementary"
<Representation id="2" width="3840" height="2160" ....>
<BaseURL>B.mp4
</BaseURL>
</Representation>
</AdaptationSet>
```

그림 3. SupplementalProperty에 Rank가 기입된 예

용자 시점을 기준으로 주변 시점의 영상 중요도가 배치되도록 한다. 영상의 중요도는 대역 할당 시에 비중을 정하는데 도움을 줄 수 있다. 중요도는 Rank # 으로 할당되며 Adaptation Set에 추가정보로 기록한다. Rank # 는 서버에 영상의 위치를 모두 파악하고 기록하고 있다는 가정 하에 측정된다. Rank # 는 콘텐츠에 따라 범위가 다양하게 바뀔 수 있다. 예를 들어 매우 빠르게 움직이는 영상을 담은 콘텐츠일 경우 피사체의 움직임에 따라 시청자의 움직임도 빨라질 수 있기 때문에 Rank # 의 범위를 늘리고, 자연풍경이나 콘서트 장 같이 움직임이 크지 않은 콘텐츠일 경우 Rank # 의 범위를 줄일 수 있다. 또한 기록된 시점이 매우 좁은 경우라면 Rank #를 더욱 세분화 하고, 기록된 시점의 거리가 매우 넓다면 Rank #를 길게 배치할 수 있을 것이다. Rank # 에 대한 정보는 Adaptation Set의 추가 정보로 기록한다.

Adaptation Set의 경우 Essential Property와 Supplemental Property 같이 추가 정보 기록을 위한 공간이 있기 때문에 이를 활용하여 그림 3과 같이 Rank # 를 기록할 수 있다. 최초 Essential Property는 이후 Adaptation Set에서 특정 기능을 활용할 것이라는 것을 명시하는 역할을 하며, 이후 Adaptation Set에는 Supplemental Property를 통해 Rank # 에 대한 정보를 기입할 수 있다. 이를 통해 사용자는 Rank 각 Adaptation Set마다 매겨진 Rank # 를 확인하여 영상 요청 시 대역의 비중 분배에 반영할 수 있다.

### 2-3. Rank 시스템을 통한 시점이동 지연 감소

2-2에서 설명한 것과 같이 시청중인 시점 혹은 디스플레이에 표출 중인 시점을 Rank #1로 배치하고 가장 높은 비중의 대역을 할당하여 높은 사용자 경험을 제공한 뒤 가장 가까운 영역의 시점들의 영상을 2순위의 대역을 할당하여 사용자가 시점이동을 할 경우 버퍼에 있는 영상을 통해 저지연의 시점 전환을 구현한다. 시점이 전환되었을 경우 현재 표출중인 영상이 변하므로 현재 시점을 중심으로 한 새로운 MPD를 전송하게 되고 이에 따라 버퍼에 저장된 영상이 소모된 뒤 새로운 MPD를 기준으로 현재 재생중인 위치의 가장 높은 순위의 영상을 다운받아 재생하게 된다. 이는 MVC와 같이 빠른 시점 변환을 제공하며, 사용자가 매우 빠른 속도로 시점을 멀리 움직이지 않는다는 가정 하에 사용자 경험 역시 매우 높은 수준으로 유지시킬 수 있다.

### 3. 제안시스템의 효과

기존 다중시점 영상 압축방법인 MVC는 주변 카메라의 이미지가 비슷한 점을 이용하여 영상의 전체 크기를 줄이는 방법을 사용한다. 하지만 3DoF 이상의 자유도를 가지는 영상의 경우 모든 방위의 영상이 기록되기 때문에 카메라를 일정거리 이상 가까이 놓을 수 없다. 일정거리 이상 가깝게 카메라를 배치할 경우 영상에 카메라가 모두 녹화되기 때문이다. 또한 MVC의 경우 땅에서 수직인 평면에서 움직이는 방향으로 카메라가 배치되지만 360 영상의 경우 땅과 수평인 평면에서 카메라가 배치되기 때문에 기존 MVC에서 고려하는 카메라 배치보다 영상의 유사도가 떨어질 확률이 매우 높다. 따라서 실질적인 MVC의 압축효율이 높지 않다고 했을 경우 시점이동의 효과 때문에 모든 시점을 하나의 파일로 압축하는 것이 더욱 비효율적일 수 있다. 따라서 제안 방식은 360 영상이 가지고 있는 특징에 맞게 다중시점 영상을 전송할 수 있는 방법을 저장하고 전송할 수 있음을 보여준다.

### 4. 결론

본 논문에서는 기존 360 영상을 기반으로 다양한 시점을 동시에 전송할 수 있는 DASH의 MPD 구조를 제시했다. 이는 적절하게 6DoF를 담아낼 수 있는 영상 표준이 없는 현재 다양한 시점을 동시에 전송 가능한 방법에 대해서 한 가지 방법을 제시했다. 현재 다양한 방법으로 공간을 영상화 하여 압축하는 방법이 표준화 진행되거나 워킹그룹으로 표준화를 위해 연구 중인 가운데 실질적인 6DoF 콘텐츠 제작 전까지 3DoF+ 혹은 6DoF인 콘텐츠를 전송하는데 있어서 중요한 연구 주제 중 하나로 자리매김 할 가능성이 있을 것으로 보인다. 따라서 제시

한 방법을 통해 시점이동을 하거나 중간 시점을 생성하여 6DoF를 전달하는 어플리케이션 및 전송 표준에 대해 연구 할 계획이다.

### ACKNOWLEDGMENT

이 논문은 2019년도 정부(과학기술정보통신부)의 재원으로 정보통신기술진흥센터의 지원을 받아 수행된 연구임. [2016-0-00144, 시청자 이동형 자유시점 360VR 실감미디어 제공을 위한 시스템 설계 및 기반기술 연구]

### 참 고 문 헌

- [1] David Nield, "How Oculus Rift works: Everything you need to know about the VR sensation", WAREABLE, 2016 (<https://www.wareable.com/oculus-rift/how-oculus-rift-works>)
- [2] MS Hololens (<https://www.microsoft.com/en-CY/hololens>)
- [3] Lafruit, Gauthier, and Siegfried Foessel. "INTERNATIONAL ORGANISATION FOR STANDARDISATION ORGANISATION INTERNATIONALE DE NORMALISATION ISO/IEC JTC1/SC29/WG1 (JPEG) & WG11 (MPEG)."
- [4] Merkle, Philipp, et al. "Efficient prediction structures for multiview video coding." IEEE Transactions on circuits and systems for video technology 17.11 (2007): 1461-1473.
- [5] Sodagar, Iraj. "The mpeg-dash standard for multimedia streaming over the internet." IEEE MultiMedia 18.4 (2011): 62-67.