

# MPEG-I Audio 기술 표준화 동향

정현주 가우디오랩(주) 연구원

## 1. 개요

ISO/IEC JTC1/SC29 WG11 126차 MPEG 회의가 2019년 3월 25일부터 29일까지 스위스 제네바에서 개최되었다. 오디오 분야에서는 MPEG-I Audio 표준화와 관련하여 129차 회의(2020년 1월)에서 제안요청서(CfP, Call for Proposal), 132차 회의(2020년 10월)에 작업 초안(WD, Working Draft), 134차 회의(2021년 4월)에 위원회 초안(CD, Committee Draft)을 발간하는 일정으로 진행하기로 계획하였다. 지난 회의에서 확정된 요구사항(Requirements)을 중심으로 MPEG-I 오디오 렌더링 기술 선정을 위해 필요한 절차를 진행 중이며, 우선적으로 오디오 장면(Audio Scene)을 표현하기 위해 필요한 모든 정보를 MPEG-I 오디오 인코더에 입력하기 위한 공통된 입력 형식(EIF, Encoder Input Format)에 대한 논의가 주를 이루었다. 또한 평가 플랫폼의 성능 개선과 기능 추가를 위한 일정을 Workplan에 업데이트하였다.

## 2. 회의 주요 결과

지난 125차 회의에서 확정된 MPEG-I Audio의 요구사항을 기준으로, 이를 평가하기 위한 입력 포맷, 절차 및 방법 등에 대한 폭넓은 논의가 이번 회의 기간에 중점적으로 진행되었다. 평가에 필요한 오디오 장면(Audio Scene)을 표현하는데 필요한 정보를 인코더에 입력하기 위한 공통된 입력 형식(EIF)을 보다 세부적으로 정의하였다. 지난 회의에 Fraunhofer-IIS를 주축으로 제안한 XML 기반 EIF 포맷을 기본 바탕으로 음향 객체의 크기 및 지향특성, 사용자와 음향 객체의 상호작용(interactivity), 반사/흡음과 같은 공간의 음향 특성 등을 표현하는 구체적인 메타데이터가 추가되었고, 음향 공간의 기하 구조를 표현하기 위한 도형, 메시 구조 및 좌표계도 확정하였다. 아래 [그림 1]은 측정된 지향성 정보를 가지는 객체가 포함된 음향 장면의 간단한 EIF 예시를 나타낸다.

```
<AudioScene>
  <AudioStream id="signal:trumpet"
    file="armstrong.wav"
    vstChannels="0, 1" />

  <SourceDirectivity id="dir:trumpet"
    file="trumpet.sofa" />

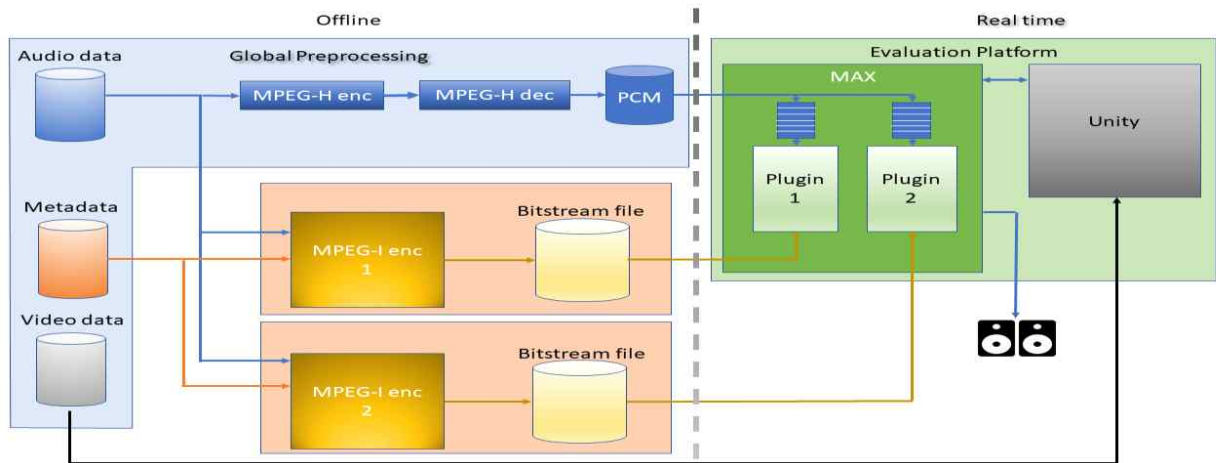
  <ObjectSource id="src:trumpet"
    position="2 1.7 -1.25"
    orientation="30 -12 0"
    signal="signal:trumpet"
    directivity="dir:trumpet"
    gainDb="-2"
    active="true" />
</AudioScene>
```

[그림 1] Audio Scene 예시

CfP의 발행 전까지 EIF의 미흡한 부분을 보완하여 포맷을 확정할 계획이며, 정의된 포맷을 활용하여 MPEG-I Audio의 렌더러 성능 평가를 위해 필요한 다양한 테스트 아이템을 선정할 계획이다. 현재까지 확보된 테스트 아이템들은 음향 장면을 구성하기 위한 EIF가 아직 정의되어 있지 않거나 MPEG-I 요구사항의 만족 여부를 평가하는데 부족한 항목이 많기 때문에 이를 보완하는 테스트 아이템의 추가가 필요한 상황이다. 추가로 확보해야 할 테스트 아이템에서 평가 항목으로 필요한 대표적인 특징은 다음과 같다.

- 음향 객체의 음상 정위 성능
- 음향 객체의 방사 패턴(지향성)
- 음향 객체의 크기
- 음향 오클루전(Occlusion)
- 공간의 반사음 등 잔향(Reverberation) 특성
- Scene Gateways(문, 창문, 복도 등)를 통한 음향 전달

확보된 테스트를 위한 음향 장면은 아래 [그림 2]와 같이 MPEG-I Evaluation Platform을 사용하여 평가를 진행한다. Evaluation Platform은 MaxMSP와 Unity를 활용하여 제작되었으며, HMD를 통해 사용자의 움직임을 트래킹하여 실시간으로 시뮬레이션/렌더링하여 음향 및 시각 정보를 사용자에게 제공한다. 공통된 오디오 신호(Audio Data)와 EIF 형태로 주어지는 메타데이터(Metadata), 그리고 Unity 기반으로 렌더링 된 비디오 데이터(Video Data)가 사용되며 제안 기술별로 각각 구현된 MPEG-I Encoder와 Plugin이 음향 렌더링을 수행한다. 플랫폼 상에서 실시간으로 제안 기술별 렌더러를 선택하여 주관 평가를 수행할 수 있다.



[그림 2] MPEG-I Evaluation Platform

MPEG-I Evaluation Platform을 사용하여 MPEG-I 제안 기술 평가를 수행하는 것에는 오디오 서브그룹에서 동의를 한 상태이지만, 보다 구체적인 평가 절차나 선정 방법에 대해서는 아직 추가 논의가 필요한 상태이다. VR 및 AR 재현을 위한 하나의 MPEG-I Audio 기술을 선정하는 것과, 실험 환경이 가능한 경우에는 VR 및 AR 평가를 모두 수행하는 것으로 합의 하였다. AR 오디오 평가에서는 HMD 대신에 Hololens와 같은 AR 헤드셋을 사용하게 되며 비디오 정보는 제한되거나 아주 간단한 형태의 위치 정보만 표시하는 형태로 평가가 진행될 예정이다. 또한 평가 항목에 대해서, 주관 평가로 진행해야 할 부분과 객관 평가로 진행해야 할 부분에 대해서 정의가 필요하며, 평가 점수를 산정하는 방식에 대해서도 많은 논의가 있었다. 보다 구체적으로는 다수의 테스트 항목과 환경을 모두 고려하여 하나의 점수(Figure of Merit)를 산정하고 이를 기반으로 최종 참조 모델(RM0)로 확정하자는 방안과 여러 개의 기술 카테고리 고리로 평가 항목을 나누어서 해당 카테고리의 최적 기술을 선정하여 이 기술들의 조합으로 최종 RM0를 확정하자는 주장이 제안되었다. 보다 구체적인 평가 방법에 대한 논의는 차기 회의에서 계속 진행될 예정이다.

### 3. 향후 전망

지난 회의에서 요구사항을 확정된 것을 계기로 6DoF 미디어 분야에 대한 MPEG-I Audio의 표준화에 대한 활발한 논의가 진행되고 있다. 평가를 위한 인코더 인풋 포맷(EIF)은 큰 무리없이 확정될 것으로 예상되지만, MPEG-I 평가 방법에 대해서는 참여 기관별로 다른 입장을 가지고 있어서, 최종 확정까지 여러 의견이 제안될 것으로 예상된다. 참여 기관별로 제안 기술의 성능 우위를 수월하게 판별할 수 있으려면 해당 기술의 장점을 부각할 수 있는 테스트 아이템의 선정이 중요한 만큼, 향후 신규 테스트 아이템도 다양하게 제안될 것으로 전망된다.