

메타버스 미디어 플랫폼과 관련 표준화 동향

□ 김상균 / 명지대학교

요약

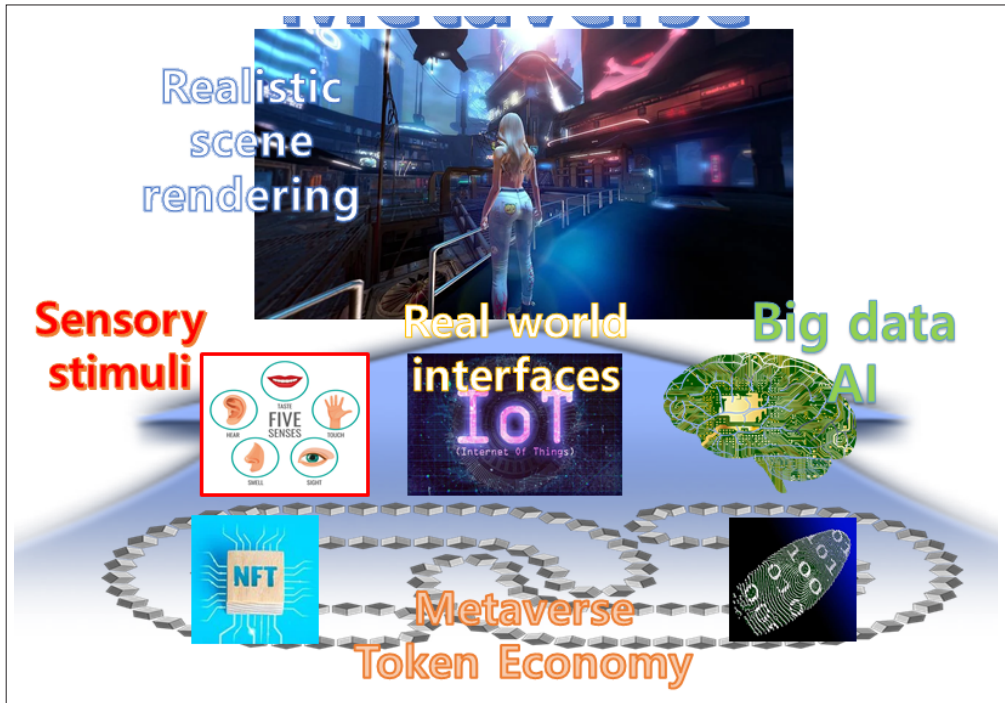
본 기고에서는 메타버스가 성공적인 미디어 플랫폼이 되기 위해서 필요한 필요 기술들을 간략하게 소개하고, 아울러 메타버스와 현실 세계 간의 원활한 데이터 교환과 제어를 위해 국제표준화를 통해 제정된 표준 기술들을 소개하고자 한다.

1. 서론

세계는 SNS를 통해 개인들이 미디어의 주체가 되어, 서로 소통하고 즐길 수 있는 새로운 플랫폼의 시대를 경험하고 있다. 유튜브를 통해 관심있는 분야의 동영상 시청하고 개인의 지식을 공유하며, 페이스북이나 인스타그램을 통해 자신과 친구들의 근황을 전하고, 틱톡을 통해 친구나 대중들에게 재미있는 짧은 동영상을 공유한다. 메타버스는 이러한 개인들의 미디어 생활에 새로운 방안을 제시한다.

메타버스 안에서 사용자들은 더 적극적인 미디어 활동을 할 수 있게 된다. 예를 들어, 자신이 만든 음악이나 비디오를 다른 사람들과 좀 더 적극적으로 공유할 수 있다. 현재의 검색이나 추천을 통한 방식에서 서로의 아바타를 통해 직접 제시되는 미디어를 감상하고 즐길 수 있게 되는 것이다. 인터넷을 통해 미디어를 생산 및 소비하는 수동적인 방식에서 사용자 주도의 좀 더 적극적인 방식으로 바뀌게 된다. 사용자는 누구나 자신이 제작한 콘텐츠를 메타버스 공간 어디서나 발표할 수 있고 이에 대한 의견을 들을 수 있으며, 아울러 콘텐츠의 양도나 매매까지도 직접 수행할 수 있다. 메타버스 내에서 즐기는 유튜브, 페이스북, 틱톡, 스포티파이, 인스타그램을 상상해 보는 것은 어려운 일이 아닐 것이다.

그러면 기존의 가상 세계(Virtual Reality)나 2000년 초반에 주창되었던 메타버스와 현재의 메타버스는 어떠한 차이가 있을까? 현재의 메타버스를 단순히 가상 공간 내에서의 게임이나 소통을 위한 서비스 또는 VR/



<그림 1> 미디어 서비스를 위한 메타버스 플랫폼 기술 구성도

AR/MR/XR 기술의 한 구현으로 보는 것은 기존의 메타버스와의 차별성을 두기 어렵다. 적어도 미디어 플랫폼으로서의 메타버스는 기술적인 측면에서 더욱 다양한 시각을 포함해야 한다.

본 기고에서는 현실 세계와 메타버스 간의 데이터 교환, 현실 객체와 메타버스 객체 간의 동기화, 오감 효과 제공을 위한 국제표준화 활동을 중심으로 새로운 메타버스에 필요한 기술들에 대한 이해를 돕고자 한다.

II. 메타버스 미디어 플랫폼의 기술 구성

<그림 1>에서 도시한 바와 같이 미디어 서비스를 위한 메타버스 플랫폼을 구성하기 위해서는 최소한 그림

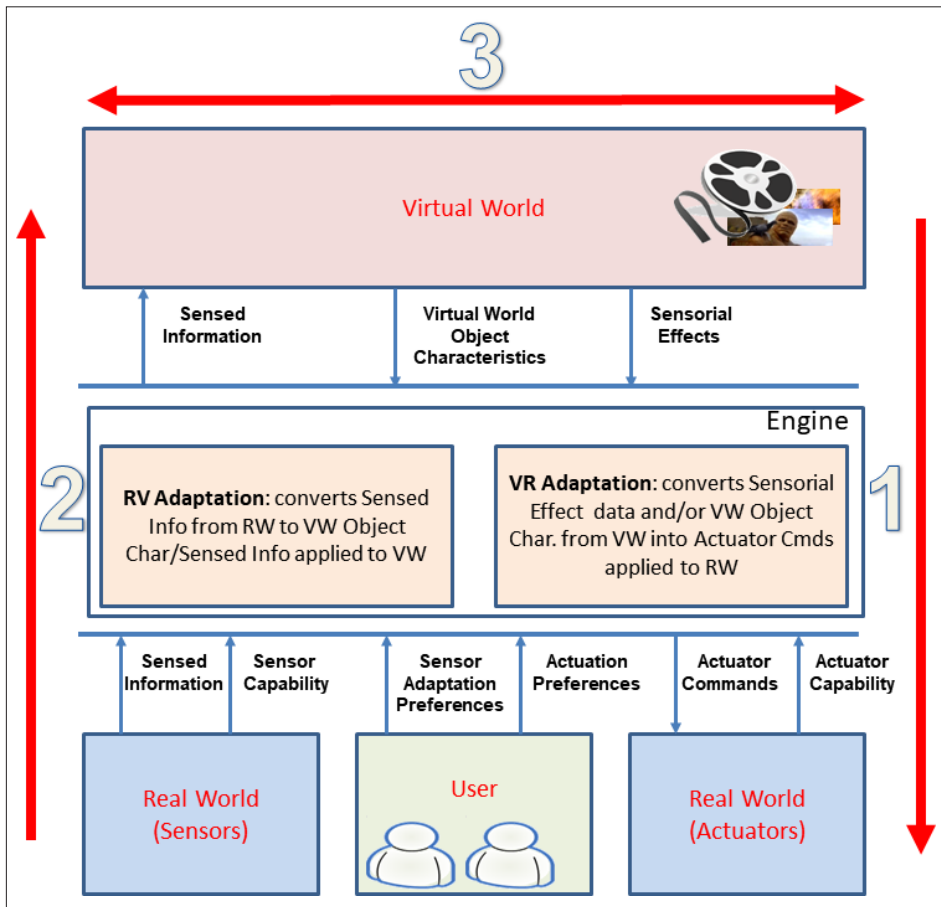
에서 제시하는 기초 기술의 융합이 필요하다. 먼저 기존 VR/AR/MR/XR 기술을 포함하여 좀 더 사실적이고 현실적인 시청각 공간 및 객체 렌더링 기술이 필요하다. 아울러 현실 세계와 메타버스 간의 정보 교류를 위한 데이터 전달 및 제어에 대한 기술들이 필요하다. 메타버스의 몰입감을 높이기 위해 실사적인 시청각 자극의 제공에 더불어 메타버스에서 경험하는 후각, 촉각, 미각 등을 현실 세계에 제공할 수 있는 오감 미디어 기술들도 중요한 역할을 할 것이다. 또한 메타버스에서 생성되는 방대한 양의 데이터를 분석하고, 다양한 솔루션을 제공할 수 있는 AI 기술과의 결합은 필수 불가결하다. 마지막으로 메타버스 내의 경제 활동을 가능케하고, 미디어 콘텐츠의 가치 및 저작권과 관련된 솔루션의 일환으로 블록체인이나 NFT와 같은 기술들의 융합이 활발해지고 있다.

III. 메타버스 미디어 플랫폼 기술 표준화 동향

1. 현실 세계와 메타버스의 조화를 위한 표준 기술

메타버스의 구현을 위해서는 현실 세계의 정보가 실시간으로 메타버스에 전달되어야 하고, 반대로 메타버스의 상황이 현실 세계에 전달되어 현실 구동기를 제어할 수 있어야 한다. 이러한 개념을 최초로 도입한 대

표적인 표준으로 ISO/IEC JTC 1 SC 29의 23005(이하 MPEG-V) 표준이 있다. MPEG-V 표준은 2000년대 초반부터 시작된 유럽의 메타버스 프로젝트와 세컨드라이프라는 가상공간에서의 제2의 생활을 주제로 한 서비스로 촉발되어 2008년부터 표준화 활동을 시작하였다. MPEG-V는 가상현실의 객체 특성 표현과 더불어 현실 세계의 컨텍스트 정보를 가상 세계에 전달하고 가상 세계(미디어 콘텐츠, 게임, 메타버스 등)의 정보를 현실 세계에 전달하여 현실 세계의 미디어 관련 구동기를 제어할 수 있는 데이터 포맷을 표준화하였다. 2018년 말



<그림 2> MPEG-V 아키텍처[1]

4차 개정판이 완료되었으며, 몰입형 미디어를 위한 추가적인 표준화가 시작될 예정이다.

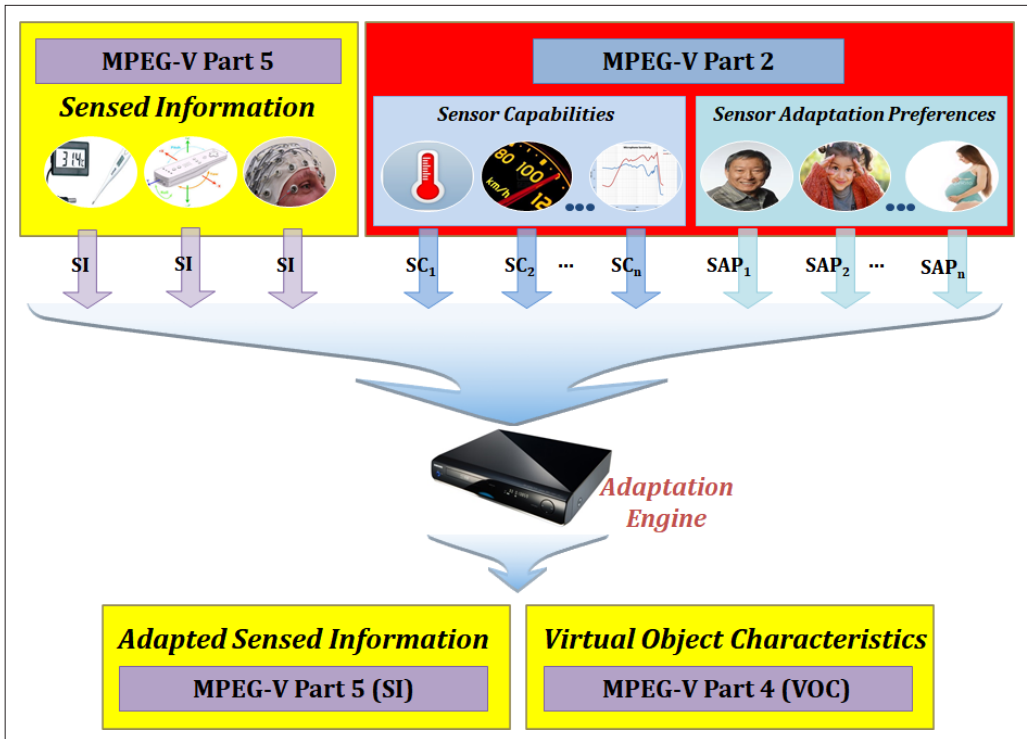
〈그림 2〉는 MPEG-V 시스템 아키텍처와 데이터 전달 시나리오의 다이어그램을 보여준다. MPEG-V 표준은 현실 세계와 가상 세계 간의 세 가지 다른 유형의 미디어 데이터 교환에 사용된다. 첫 번째 미디어 교환은 가상 세계에서 현실 세계로의 정보 전달 및 데이터 적응이다(〈그림 2(1)〉). 감각 효과 데이터(MPEG-V 3부) 및 가상 세계 개체 특성(MPEG-V 4부)을 컨텍스트 입력으로 받아들인다. MPEG-V 2부에 정의된 구동기 성능(Actuator Capability), 구동 선호도(Actuation Preferences) 및 MPEG-V 5부에 정의된 센서 데이터 정보(Sensed Information)를 제어 매개변수로 사용한다. 이를 통해 현실 세계의 구동기에 전달되는 구동기 명령(Actuation Command)을 MPEG-V 5부의 데이터 포맷을 이용하여 생성한다. 가상-2-현실(Virtual-2-Real: V.R.) 적응 엔진은 입력 제어 매개변수에 따라 가상 세계의 객체 특성 또는 가상 세계의 감각 효과 데이터를 현실 세계의 구동기 명령으로 변환(적용)한다.

두 번째 미디어 교환은 현실 세계에서 가상 세계로의 정보 전달 및 적응이다. 실제 상황에서 센서의 감지 정보(MPEG-V 5부)를 표준화된 방식으로 생성한다. MPEG-V 2부에 정의된 센서 성능(Sensor Capability), 센서 적응 선호도(Sensor Adaptation Preferences)를 제어 매개변수로 사용한다. 이를 통해 MPEG-V 4부에 정의된 가상 세계 개체 특성(Virtual Object Characteristic)과 가상 세계를 위해 적응된 감지 정보(Adapted Sensed Information)를 생성한다(〈그림 2(2)〉). 현실-2-가상(Real-2-Virtual: R.V.) 적응 엔진은 입력 제어 매개변수에 따라 현실 세계의 센서로부터 감지된 정보를 가상 세계 객체 특성 및 가상 세계에 적응된 감지 정보로 변환(또는 조정)한다.

마지막으로 가상 세계 간의 정보 교환은 고유한 가상

세계 객체 특성을 MPEG-V 4부에 규범적으로 지정된 가상 세계 객체 특성으로 표현하여 수행한다(〈그림 2(3)〉). 이는 메타버스를 구축하고 운영하는 기업이 늘어날수록 더 중요한 의미를 가질 수 있다. 메타버스의 구축은 어느 기업이나 가능한 상태에서 세상에는 다수의 메타버스(멀티버스)가 상용화될 수 있고, 사용자는 이러한 메타버스 간에 자신의 디지털 자산(Digital Assets)을 손쉽게 이동할 수 있는 방법이 필요하게 된다. 이때 하나의 메타버스 내 디지털 자산을 표준화된 방법으로서 술하게 되면 이는 다른 메타버스에서도 이를 해석할 수 있어 각자의 방식에 맞는 디지털 자산으로 변환 가능하게 된다. MPEG-V 4부의 가상 세계 객체 특성(Virtual World Object Characteristics) 표준은 이를 가능하게 해주는 가상 세계 객체나 아바타의 표현에 대한 표준을 정의한다.

〈그림 3〉은 MPEG-V의 각 프로젝트가 제공하는 데이터를 이용하여 현실 세계의 컨텍스트 정보를 가상 세계(예: 메타버스)로 전달하는 데이터 흐름을 예시한다. 현실 세계의 센서로부터 측정되는 데이터는 MPEG-V 5부에서 정의된 센서(Sensed Information) 데이터 포맷으로 표현되어 가상 세계에 전달된다. 이때 정밀한 데이터 표현을 위해서는 현실 세계 센서의 정확한 성능에 대한 표현이 필요한데, 이때 MPEG-V 2부에서 정의된 센서 성능(Sensor Capability) 데이터 포맷을 이용하게 된다. 가상 세계는 입력된 센서 데이터와 센서의 성능 정보를 바탕으로 가상 세계 렌더링에 적합한 적응된 센서 데이터(Adapted Sensed Information)로 변환하여 사용할 수 있다. 아울러 센서 데이터에 대하여 사용자가 일종의 제약을 둘 수도 있는데 이를 센서 데이터에 대한 사용자 선호도(User Preference)로 표현할 수 있다. 센서 데이터와 센서에 대한 사용자의 선호도를 결합하여 적응된 센서 데이터를 생성할 수 있으며, 이는 다시 가상 세계의 객체(Virtual Objects)를 표현하거나 가상 세



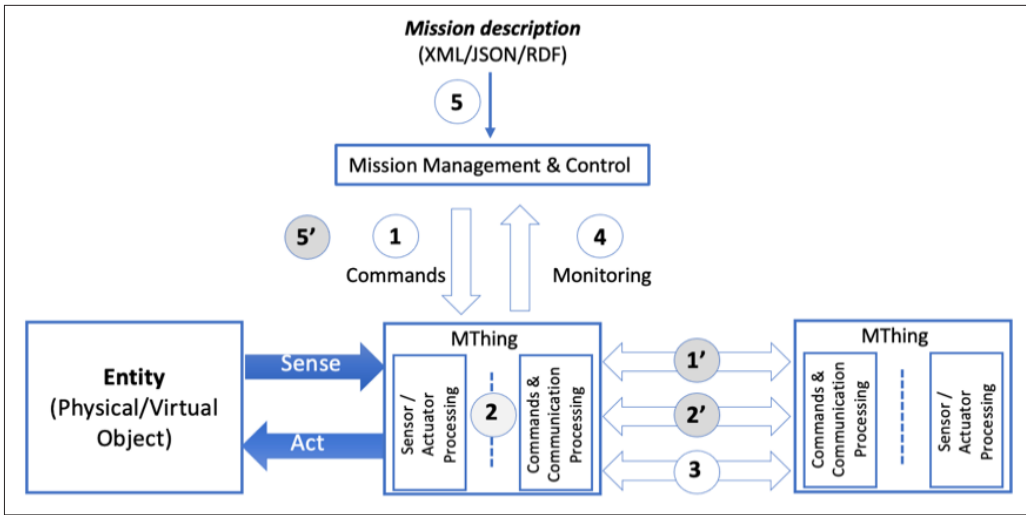
<그림 3> 현실 세계의 센서 데이터 정보, 센서의 성능 정보 및 사용자 선호 정보를 결합하여 가상 세계 센서 데이터 정보로 변환한 예

계의 NPC를 운영하는 데 활용할 수 있다.

MPEG-V가 미디어 서비스를 중심으로 한 현실 세계와 가상 세계 간의 데이터 포맷을 표준화했다면, 현실 세계와 가상 세계의 연결된 사물을 중심으로 자율적인 미디어 서비스를 제공하기 위한 국제표준이 진행되고 있다. 바로 ISO/IEC JTC 1 SC 29 그룹에서 진행 중인 23093(이하 MPEG-IoMT: 미디어 사물인터넷) 국제표준이다. 2021년 7월 현재 개정판(2nd edition)에 대한 FDIS 단계가 진행 중이다. 이는 현실 세계의 미디어와 관련된 사물들을 미디어 사물(Media Things)로 정의하고, 미디어 사물을 미디어 센서(MSensor), 미디어 분석기(MAnalyser), 미디어 저장소(MStorage), 미디어 구동기(MActuator) 등으로 세분화하여 이들 사물 간에 교

환하는 데이터와 데이터 교환을 위한 API를 표준화한다. 특히 복수 개의 미디어 사물들이 자율적으로 연동하여 운영될 수 있도록 사용자가 정의하는 미디어 서비스를 표준화된 임무(Mission or Task) 데이터로 표현하는 방식에 대하여 새로운 프로젝트(ISO/IEC 23093-5)를 시작할 예정이며, 메타버스와 같은 좀 더 광의의 서비스를 위한 미디어 사물 데이터 및 API를 포함하는 3차 개정판(3rd Edition)에 대한 표준화를 2022년 초반부터 시작할 예정이다.

<그림 4>는 MPEG-IoMT 아키텍처를 보여준다. 제시된 아키텍처는 현재 또는 새로운 IoMT의 확장에 속할 인터페이스, 프로토콜 및 미디어 관련 정보 표현을 제시한다. <그림 4>의 IoMT 아키텍처는 아래와 같이 인터페



<그림 4> MPEG-IoMT 아키텍처[3]

이스, 프로토콜 및 미디어 관련 정보 표현을 정의한다.

- 인터페이스 1 : 시스템 관리자와 미디어 사물 간의 사용자 명령(설정 정보)
- 인터페이스 1' : 인터페이스 1의 수정된 형식(예: 인터페이스 1의 서브 세트)으로 미디어 사물에 의해 다른 미디어 사물로 전달되는 사용자 명령(설정 정보)
- 인터페이스 2 : 미디어 사물에 의해 감지된 데이터 (원시 또는 처리된 데이터) (압축 또는 의미 추출) 및 구동(actuation) 정보
- 인터페이스 2' : 변환된 인터페이스 2(예: 전송용)
- 인터페이스 3 : 미디어 사물의 특성(Characteristics) 및 발견(Discovery)
- 인터페이스 4 : 미디어 사물의 수행 상태 모니터링 정보
- 인터페이스 5 : 사용자가 IoMT 네트워크에 할당된 미션을 표현하는 구조화된 데이터 포맷(예: XML/JSON/RDF)

- 인터페이스 5' : 임무의 관리나 제어를 위해 하나 또는 여러 미디어 사물로 전달되는 구조화된 데이터 포맷(예: XML/JSON/RDF), 인터페이스 5의 수정된 형식(예: 인터페이스 5의 하위 집합)

이 외에도 2020년 10월 IEEE 2888.3에서 시작된 사이버 세계와 물리적 세계 간의 디지털 동기화 조정에 관한 표준(Standard on Orchestration of Digital Synchronization between Cyber and Physical World)이 있다. 이 표준은 CPS(Cyber-Physical System), DTS(Digital Twin System), 메타버스를 구현하려는 연구원 및 산업 종사자에게 표준화된 지침을 제공하기 위한 것이다. 물리적 개체의 동기화 및 상호 작용 시퀀스를 제공하기 위해 디지털 개체에 대한 매개변수를 설정하고 디지털 개체와 통신하기 위한 어휘, 요구 사항, 평가 방법, 데이터 형식 및 API(응용 프로그램 인터페이스)를 정의한다.

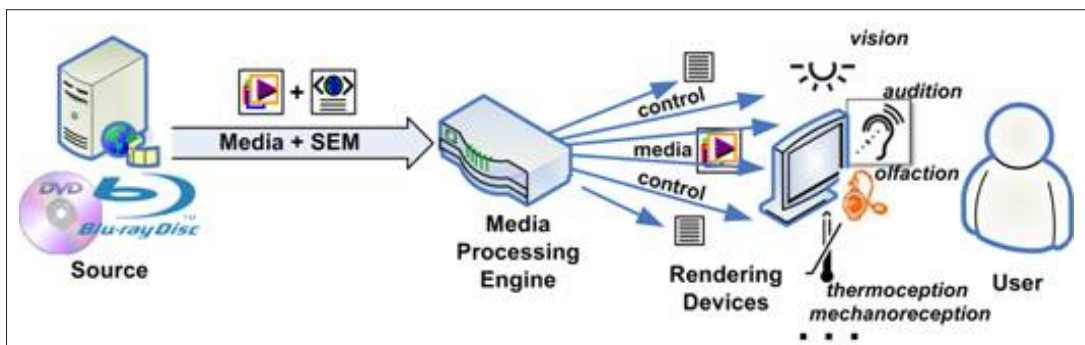
2. 현실 세계의 오감 만족을 위한 표준 기술

메타버스 세계에서 경험할 수 있는 시각, 청각 이외의 후각(Olfactory), 촉각(Tactile) 등의 오감 효과를 현실에서도 경험하기 위해서는 감각 효과(Sensory Effect)를 표현할 수 있는 표준이 필요하다. 메타버스 내에서 발생하는 후각이나 촉각 효과는 감각 효과 메타데이터로 표현되어 현실 세계에 실시간으로 전달되고, 이는 다시 사용자의 효과에 대한 선호도(User Preference)와 효과를 렌더링하는 구동기의 성능(Actuator Capability)을 고려하여 최종 감각 효과 명령어(Sensory Command)로 변환된다.

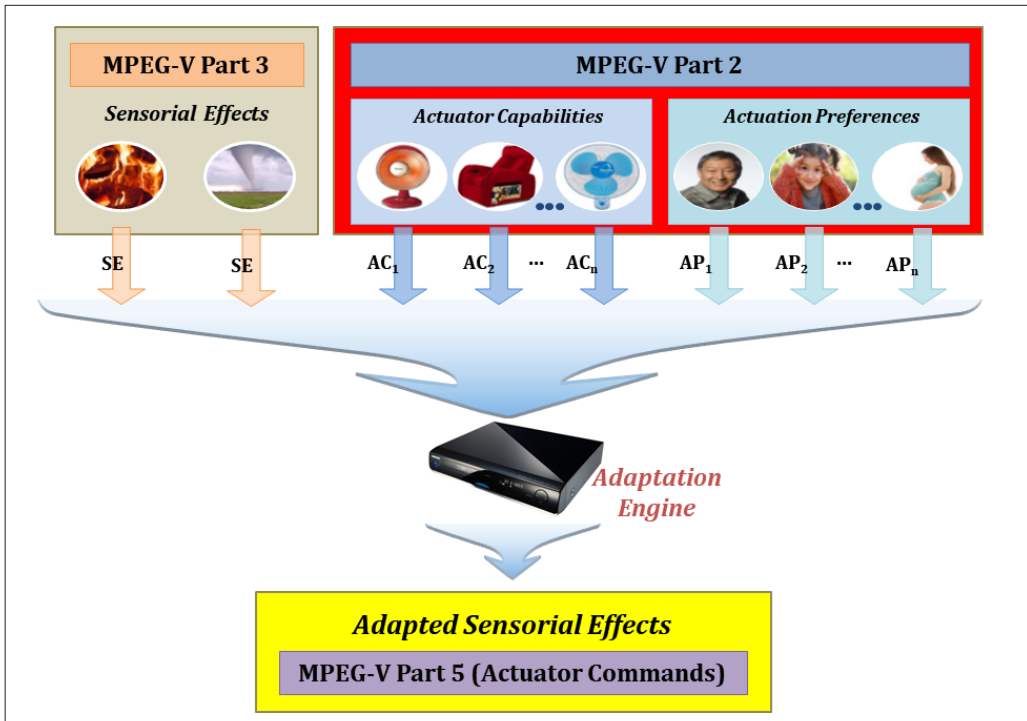
MPEG-V 3부(ISO/IEC 23005-3) 감각 정보(Sensory information)는 이러한 감각 효과를 표현하기 위해 감각 효과 서술 언어(SEDL: Sensory Effect Description Language)[2]를 XML 스키마 기반의 언어로 정의한다. 이는 인간의 감각을 자극하는 빛, 바람, 안개, 진동, 온도, 냄새 등과 같은 감각 효과를 표현할 수 있게 해준다. 실제 감각 효과는 확장성과 유연성을 위해 별도의 감각 효과 어휘(SEV: Sensory Effect Vocabulary) 내에서 정의되어 각 응용 프로그램 영역에서 감각 효과를 표현할 수 있다. 감각 효과 서술 언어(SEDL)의 문법을 준

수하는 서술 예를 감각 효과 메타데이터(SEM: Sensory Effect Metadata)라 하며 모든 멀티미디어 콘텐츠(예: 영화, 음악, 웹사이트, 게임, 메타버스)는 이를 통해 콘텐츠 내의 감각 효과를 표현할 수 있다. 감각 효과 메타데이터(SEM)는 팬, 진동 의자, 향기 발생기, 램프 등과 같은 구동기를 적절한 중개 장치(미디어 처리 엔진)를 통해 조종하여 콘텐츠 내 오감 효과에 대한 현실 세계의 사용자 경험을 높이는 데 사용된다. 예를 들어 가상공간이나 영화 콘텐츠 내 시청각 효과와 더불어 사용자는 후각, 촉각 효과를 감지할 수 있어, 사용자에게 특정 미디어의 일부인 느낌을 줌으로써 좀 더 가치 있고 유익한 사용자 경험을 제공할 수 있다.

<그림 5>는 시청각 정보 외에 감각 효과를 제공하는 개념에 관해 설명한다. 메타버스와 같은 가상공간에서의 오감 효과는 실시간으로 감각 효과 메타데이터(SEM)로 변환되어 미디어 처리 엔진에 전달될 수 있다. 이때 미디어 처리 엔진은 중개 장치 역할을 하며, 미디어 및 감각 효과 렌더링에 대한 사용자 설정(User Preference)을 기반으로 동기화된 방식으로 실제 미디어 리소스 및 수반되는 감각 효과를 재생하는 역할을 한다. 아울러 미디어 처리 엔진은 다양한 렌더링 장치의 성능(Actuator Capability)에 따라 미디어 리소스와 감



<그림 5> MPEG-V 감각 효과 서술 언어의 개념[4]



<그림 6> 가상공간의 감각 효과와 현실 세계의 구동기 성능 및 사용자의 구동 선호도를 결합하여 생성된 적응된 감각 효과[1]

각 효과 메타데이터(SEM)를 모두 조정할 수 있다.

메타버스에서 전달되는 감각 효과를 현실 세계의 감각 효과 명령어(Sensory Command)로 변환하기 위해서는 MPEG-V 2부(ISO/IEC 23005-2)의 현실 세계의 구동기 성능(Actuator Capability) 및 구동기에 대한 사용자의 선호도(User Preference)를 표현하는 서술 인터페이스가 필요하다. 구동기에 대한 사용자의 선호도는 적응 엔진에 추가 정보를 제공하여 실제 세계 제어를 위한 구동 명령(Actuation Command)의 미세 조정에 사용할 수 있다. 이러한 감각 효과 구동 명령은 MPEG-V 5부(ISO/IEC 23005-5)에서 정의된 감각 효과 명령어(Sensory Command) 서술 인터페이스를 통해 표현된다.

<그림 6>과 같이 감각 효과(Sensory Effects)는 구

동기 명령어(Actuation Command) 및 구동 선호도(Actuation Preference)에 따라 적응된 감각 효과(Adapted Sensory Effects)로 조정될 수 있다. 예컨대 가상공간 내의 바람의 세기(감각 효과의 세기)가 토네이도급의 강풍을 표현하더라도 사용자가 가지고 있는 팬의 성능에 따라 팬이 가지고 있는 가장 높은 바람 세기로 표현될 수 있다. 단, 이는 사용자에게 불편함을 초래할 수 있으므로 사용자는 최대 바람 세기나 바람 지속 시간 등에 제한(구동 선호도)을 둘 수 있다. 미디어 적응 엔진은 이러한 정보를 종합하여 사용자의 구동기 성능과 감각 효과 선호도를 반영한 적응된 감각 효과 명령어(Adapted Sensory Command)를 생성하여 현실 세계의 구동기를 제어한다.

IV. 결론

본 기고는 미디어 플랫폼으로의 메타버스의 가능성을 고찰하고자 현존하는 주요 기술들이 메타버스를 구성하는 데 어떠한 역할을 할 수 있는지 간단하게 고찰

하였다. 아울러 제시된 주요 기술 중 기존 국제표준화를 통해 제시되었던 표준 기술들을 소개하였다. 메타버스가 국제표준을 차용한다면 더욱 확장성 있게 다양한 미디어 서비스나 플랫폼으로의 역할을 수행할 수 있을 것으로 기대한다.

참고 문헌

- [1] J.J. Han, S.-K. Kim, Text of white paper on MPEG-V, ISO/IEC JTC 1/SC 29/WG 11/N14187
- [2] S.-K. Kim (eds.), Text of ISO/IEC FDIS 23005-3 4th edition Sensory Information, ISO/IEC JTC 1/SC 29/WG 11/N17368, ISO/IEC JTC 1 SC 29 WG 11(2018), Gwangju, S. Korea
- [3] Technologies under consideration on MPEG-IoMT, ISO/IEC JTC 1/SC 29/WG 07/N00110 (2021)
- [4] M. Walti, C. Timmerer, H. Hellwagner, A Test-Bed for Quality of Multimedia Experience Evaluation of Sensory Effects, Proceedings of the First International Workshop on Quality of Multimedia Experience (QoMEX 2009), San Diego, USA

필자 소개



김상균

- 1991년 : 아이오와 대학(U of Iowa) 전산과학 학사
- 1995년 : 아이오와 대학(U of Iowa) 전산과학 석사
- 1997년 : 아이오와 대학(U of Iowa) 전산과학 박사
- 1997년 ~ 2007년 : 삼성종합기술원 전문연구원
- 2007년 ~ 현재 : 명지대학교 융합소프트웨어학부 교수
- 주관심분야 : 미디어 메타데이터 국제표준, 오감미디어, 미디어 사물인터넷, 블록체인, 메타버스