

Multi-Screen Service Forum Specification

MSS.S-Y19-003

서 문

3DoF+ 영상 처리 및 전송을 위한 몰입형
미디어 표준에 대한 요구사항 정의
Requirement Definition of Immersive Media
for 3DoF+ Video Processing and Streaming

제출일 : 2019년 10월 6일
제출기관 : 멀티스크린서비스포럼
제출인 : 류은석

1. 표준의 목적

이 표준의 목적은 최근에 급격히 발전하고 있는 몰입형 미디어를 위한 가상 현실 (virtual reality, VR) 기술 중 하나인 3DoF+의 구현을 위해 여러 위치에서 촬영된 다수의 360 영상들을 압축 및 전송하고 사용자의 머리 움직임에 대응하는 영상을 렌더링할 때 필요한 요구사항을 정의함에 있다.

2. 주요 내용 요약

이 표준은 사용자 움직임 추적이 가능한 머리장착형 영상장치(head-mounted display, HMD)를 통한 다수의 360도 영상들을 전송 및 렌더링하여 3DoF+를 위한 서비스를 제공할 때, 제한된 대역폭과 연산량을 사용하여 해당 서비스를 구현하기 위한 요구사항들을 기술한다.

3. 인용 표준과의 비교

해당 사항 없음.

Preface

1 Purpose

The purpose of this standard is to define the requirements of implementation on 3DoF+, which is one of recently growing virtual reality(VR) technologies for immersive media, and it requires multiple 360 videos compression, streaming, and a viewport rendering which corresponds to the user's head movements.

2 Summary

The standard describes the requirements of 3DoF+ service with a head-mounted display(HMD) which traces the user's head movement with a limited resources such as bandwidth and computing power, although it requires multiple 360 videos to be transmitted and rendered simultaneously.

3 Relationship to Reference Standards

- 1 적용 범위 4
- 2 인용 표준 4
- 3 용어 정의 4
- 4 약어 4
- 5 3DoF+ 영상 처리 및 전송을 위한 몰입형 미디어 표준에 대한 요구사항 정의 5
- 5.1 제한된 연산량 및 대역폭에서의 다수의 360 비디오 압축 및 전송을 위한 요구사항 정의 5

목 차

3DoF+ 영상 처리 및 전송을 위한 몰입형 미디어 표준에 대한 요구사항 정의

1. 적용범위

본 표준의 적용 범위는 비디오 통신에서의 시그널링을 처리하는 객체를 다루며, 이는 사용자 단말, 서버, 중계 시스템 및 라우터 등을 포함한다. 또한, 본 표준의 시그널링 메시지는 (1) Session 정보를 실어나르는 high-level syntax 프로토콜을 통해 전해질 수도 있고, (2) 비디오 표준의 SEI, VUI, 또는 슬라이스 헤더 (Slice Header) 등의 패킷 단위에서 전해질 수도 있고, (3) 비디오 파일을 설명(Descript)하는 별도의 파일로(예: DASH의 MPD) 전달될 수 있다.

2. 인용 표준

"H.265: High efficiency video coding". ITU. 2015-07-09. Retrieved 2015-08-02.

3. 용어 정의

해당사항 없음

4. 약어

AVC	Advanced Video Coding
DoF	Degrees of Freedom
HEVC	High Efficiency Video Coding
HTTP	Hyper Text Transfer Protocol
JCTVC	The Joint Collaborative Team on Video Coding
MPD	Media Presentation Description
MPEG	Moving Picture Experts Group
MV-HEVC	Multi-View HEVC
OMAF	Omnidirectional Media Format
VVC	Versatile Video Coding
SEI	Supplemental Enhancement Information

5. 3DoF+ 영상 처리 및 전송을 위한 몰입형 미디어 표준에 대한 요구사항 정의

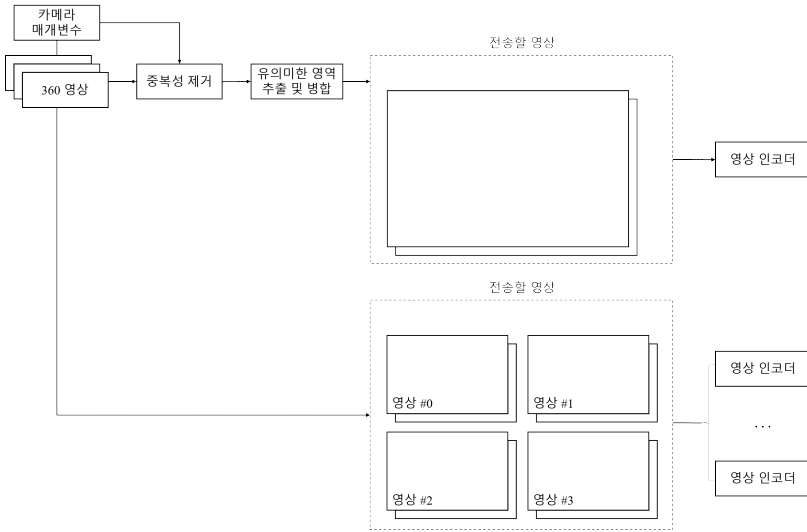
5.1. 제한된 연산량 및 대역폭에서의 다수의 360 비디오 압축 및 전송을 위한 요구사항 정의

최근 VR 기술 및 장비의 발달과 함께 여러 기업을 통해 HMD와 같은 착용 가능한 기기들이 선보이고 있다. 이를 통한 여러 서비스 시나리오 중에는 가장 대표적인 영화 관람 및 게임뿐만 아니라 화상회의와 원격 수술 등이 있다. 이를 위해서는 사용자의 머리 움직임에 대응하여 위화감이 없는 영상을 제공할 수 있어야 하는데, 해당 서비스의 예시로 3DoF+와 6DoF가 있다. 6DoF는 머리장착형 영상장치를 착용한 사용자가 걸어다니는 상황을 지원하고, 3DoF+는 의자에 앉아 있는 사용자가 머리를 움직였을 때의 상황을 대응한다. 사용자의 모든 머리 움직임을 예측하여 영상을 취득하는 것은 불가능에 가깝기에 이를 위한 시스템은 기존에 촬영된 다수의 360 영상들을 기반으로 머리 장착형 영상 장치에서 사용자 시점에 대응하는 가상 영상을 생성하기 위해 압축된 해당 영상들을 전송받아 해당 시점에 렌더링할 수 있어야 한다.

사용자에게 다수의 360 영상들을 제공하기 위해서는 해당 영상들을 취득 후 압축할 수 있어야 하는데, 일반적으로 사용되는 AVC, HEVC의 경우 영상 하나 당 하나의 인코더가 필요하다. 이를 그대로 이용할 경우 영상 취득 장치가 동시에 구동해야 할 인코더의 개수가 늘어나게 되고 이는 제한된 연산량을 사용하여 다수의 영상들을 압축하기 어렵게 만든다. 때문에 3DoF+ 영상 처리 및 전송을 위하여 인코더의 개수는 영상의 개수보다 적어야 한다. 또한 영상의 개수가 늘어날수록 전송해야 할 데이터의 양도 많아지는데, 기존 코덱을 사용할 경우 대역폭 제한을 초과하는 데이터가 생성될 확률이 높다. 일반적으로 동시에 취득된 다수의 360 영상들 간에는 중복된 부분이 존재하기에, 이를 제거하여 대역폭을 절약할 수 있다. 따라서, 대역폭을 준수할 만한 영상 간 중복성 제거 기법이 필요하다.

위의 문제를 해결하기 위해 MV-HEVC 등의 확장 표준이 등장하였으나, 하드웨어 가속을 지원받기 위해서는 해당 표준을 이용한 하드웨어가 장치에 내장되어야 한다. 하지만 이 경우 영상 압축을 위한 여러 하드웨어가 난립하게 되고, 이는 표준화의 취지에 어긋난다. 따라서 해당 서비스를 위해서는 인코더에 독립적인 시스템이 필요하다.

(그림 5-1)은 전송된 세 가지의 요구사항을 반영한 3DoF+ 시스템의 구조도를 나타낸다. 카메라의 위치, 방향 등을 나타내는 카메라 매개변수를 이용하여 다수의 360 영상들 간 중복성을 제거하여 대역폭을 절약할 수 있다. 이어, 중복성이 제거되지 않은 영역, 다시 말해 유의미한 영역들만을 추출하여 하나의 영상으로 병합하는 과정을 거치면 영상의 개수가 줄어들고, 단일 인코더/디코더 만으로 영상 부호화, 전송, 복호화가 가능하다.



(그림 5-1) 3DoF+ 영상 처리 및 전송을 위한 요구사항을 반영한 시스템 구조도