

정보통신단체표준(국문표준)  
TTAx.xx-xx.xxxx/R1

제정일: 200x 년 xx 월 xx 일  
개정일: 200x 년 xx 월 xx 일

# TTA Standard

(기술보고서는 TTA Technical Report)

머리장착형 영상장치를 위한 사용자  
시선 이동속도 기반 360 비디오 영  
상품질 제어 시그널링

360 Video Quality Control Signaling on HMD  
Considering Viewport Movement

표준초안 검토 위원회	디지털콘텐츠 프로젝트그룹(PG610)				
표준안 심의 위원회	소프트웨어/콘텐츠 기술위원회(TC6)				
	성명	소속	직위	위원회 및 직위	표준번호
표준(과제) 제안	류은석	가천대학교	조교수	-	
표준 초안 작성자	류영일	가천대학교	대학원생	-	
	손장우	가천대학교	대학원생	-	
	류은석	가천대학교	조교수	-	
사무국 담당					

본 문서에 대한 저작권은 TTA에 있으며, TTA와 사전 협의 없이 이 문서의 전체 또는 일부를 상업적 목적으로 복제 또는 배포해서는 안 됩니다.

본 표준 발간 이전에 접수된 지식재산권 협약서 정보는 본 표준의 '부록(지식재산권 협약서 정보)'에 명시하고 있으며, 이후 접수된 지식재산권 협약서는 TTA 웹사이트에서 확인할 수 있습니다.

본 표준과 관련하여 접수된 협약서 외의 지식재산권이 존재할 수 있습니다.

발행인 : 한국정보통신기술협회 회장

발행처 : 한국정보통신기술협회

13591, 경기도 성남시 분당구 분당로 47

Tel : 031-724-0114, Fax : 031-724-0109

발행일 : 20xx.xx

# 서 문

## 1 표준의 목적

이 표준의 목적은 최근 부상하고 있는 가상 현실(Virtual Reality) 기술과 머리장착형 영상장치(Head-Mounted Display; HMD)를 활용한 실시간 영상 전송 시스템에서의 사용자 시선 이동속도 정보 기반 비디오 품질 제어 정보를 시그널링(Signaling)함으로써, 불필요한 영상 전송 대역폭을 낮추고, 원활한 영상 통신이 가능하도록 하는 기술을 설명함에 있다.

## 2 주요 내용 요약

이 표준은 머리장착형 영상장치를 통한 실시간 영상 전송 관련 서비스를 제공할 때, 사용자의 시선 이동정보에 기반한 영상품질 제어정보를 시그널링 하는 기술을 담고 있으며, 기술의 장점을 설명하기 위해 (1) 본 표준의 필요성 및 스케일러블 비디오 및 머리장착형 영상장치의 배경지식, (2)본 표준을 머리장착형 영상장치를 활용한 실시간 영상 전송 시스템에 적용했을 때 기대되는 성능향상을 설명한다. 그리고, (3) 제안하는 표준 시그널링 규격(구문 (신텍스), 의미론(시멘텍스))을 기술한다.

## 3 인용 표준과의 비교

### 3.1 인용 표준과의 관련성

이 표준은 국제 표준단체 The Joint Collaborative Team on Video Coding (JCT-VC)의 Scalable Extension of High Efficiency Video Coding (SHVC) 기반 기술 등을 이용하는 시스템을 위한 별도의 독립적인 시그널링 표준으로써, SHVC 표준과 직접적인 관련성이 없음.

## Preface

### 1 Purpose

The purpose of this standard is to define the signaling syntax and semantics of the user sight movement speed-based video quality control information for HMD. The signaling standard can decrease the bandwidth requirements of video transmission and communicate seamlessly.

### 2 Summary

The standard includes the specifications for the user sight movement information signaling for VR service using HMD. It consists of three sections: (1) the motivation of this standard with the background of scalable video and HMD, (2) an expecting performance gains from applying the standard to HMD and 360 video streaming systems, and (3) the standard signaling specifications including syntax and semantics.

### 3 Relationship to Reference Standards

The standard can use the referenced video coding standard specifications such as the Scalable High Efficiency Video Coding (SHVC) of the Joint Collaborative Team on Video Coding (JCT-VC). But, this standard does not directly affect to or influenced by the referenced standard but specifies the signaling details independently.

## 목 차

1 적용 범위 .....	1
2 인용 표준 .....	1
3 용어 정의 .....	1
4 약어 .....	1
5 머리장착형 영상장치를 위한 사용자 시선 이동속도 기반 360 비디오 영상품질 제어 시그널링 .....	2
5.1 본 표준의 필요성 .....	2
5.2 본 표준 기술을 통해 기대되는 성능향상 .....	3
5.3 제안하는 표준 시그널링 규격 .....	5
5.4 추가 확장적 사용 .....	8
부록 I-1 지식재산권 요약서 정보 .....	10
I-2 시험인증 관련 사항 .....	11
I-3 본 표준의 연계(family) 표준 .....	12
I-4 참고 문헌 .....	13
I-5 영문표준 해설서 .....	14
I-6 표준의 이력 .....	15

# 머리장착형 영상장치를 위한 사용자 시선 이동속도 기반 360 비디오 영상품질 제어 시그널링

## 1 적용 범위

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) 세션(Session) 정보를 실어나르는 고수준 구문(high-level syntax) 프로토콜을 통해 전해질 수도 있고, (2) 360도 비디오 부호화 표준의 부가 정보를 전달하는 Omnidirectional Media Application Format (OMAF)를 통한 시그널링으로 전해질 수도 있으며, (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

해당 사항 없음

## 4 약어

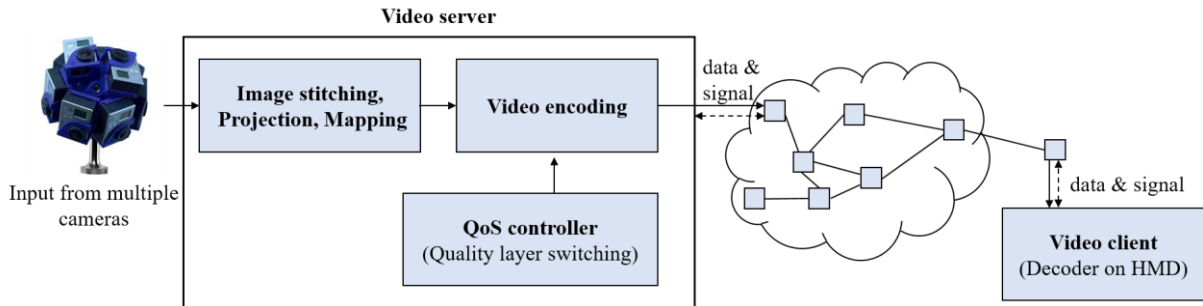
- BL: Base Layer
- EL: Enhancement Layer
- EC: Error Concealment
- HEVC: High Efficiency Video Coding
- SHVC: Scalable High Efficiency Video Coding
- UHD: Ultra High Definition
- MPEG: Moving Picture Experts Group
- JCTVC: The Joint Collaborative Team on Video Coding
- SEI: Supplemental Enhancement Information
- VUI: Video Usability Information
- DASH: Dynamic Adaptive Streaming over HTTP
- MPD: Media Presentation Description
- CU: Coding Unit
- OMAF: Omnidirectional Media Application Format

## 5 머리장착형 영상장치를 위한 사용자 시선 이동속도 기반 영상품질 제어정보 시그널링

본 섹션에서는 머리장착형 영상장치(HMD)를 위한 사용자 시선 이동속도 기반 영상품질 제어정보 시그널링 표준에 관하여 3가지를 구체적으로 설명한다: (1)필요성, (2)기대되는 성능향상, (3)구체적인 시그널링 규격.

### 5.1 본 표준의 필요성

최근 가상 현실 기술 및 장비의 발달과 함께 여러 기업을 통해 머리장착형 영상장치와 같은 착용 가능한 기기들이 선보이고 있다. 이를 통한 여러 서비스 시나리오 중에는 가장 대표적으로 실시간 360도 영상 전송 서비스가 있다.



(그림 5-1) HMD를 위한 360도 영상전송 시스템의 개념적 구성

360도 영상 전송 시스템은 그림 5-1과 같이 다수의 카메라를 사용하여 360도 영상을 취득하고, 입력된 영상을 부호화 하여 사용자가 착용하고 있는 머리장착형 영상장치로 전송한다. 전송된 영상은 복호화 후, 360도 가상공간에 맵핑(Mapping)되어 사용자에게 제공되게 된다. 이 때, HMD가 사용자의 눈과 매우 가까운 위치에서 영상을 재생하기 때문에 이질감 없는 몰입감을 사용자에게 제공하기 위해서는 UHD 급 이상의 영상을 이용할 필요성이 있는데, 이 경우 증가된 비디오 데이터 양으로 인하여 사용자 단말 및 영상전송 시스템의 대역폭 확보가 어려운 문제점이 발생하게 된다. 이러한 문제점을 해결하기 위하여 최근 360도 영상전송 시스템의 경우, 360도 영상을 서브영역으로 분할하여 특정 서브영역들만 선택적으로 전송하는 타일링 기술과 스케일러블 영상 부호화 기술을 활용하여 사용자가 바라보는 부분의 영역만 고품질 영상 정보를 전송하고, 나머지 영역은 기본적인 영상정보만 전송하여 사용자의 시선변화에 따른 화면전환속도 향상 및 영상전송 요구 대역폭을 절감할 수 있다. 하지만, 사용자의 시선변화가 빠르고 자주 일어나면, 전송해야 하는 고품질 영상 정보가 증가하여 영상전송 요구 대역폭 절감효과가 저하되게 되어 타일링 기술과 스케일러블 영상 부호화 기술을 활용하는 의미가 없어지게 된다. 본 표준은 이의 해결책으로 ‘사용자 시선 이동속도 기반 영상품질 조절정보 시그널링’ 기법의 내용과 장점, 그리고 이를 통한 대역폭 절감효과 향상을 설명한다.

### 5.1.1. 배경 지식

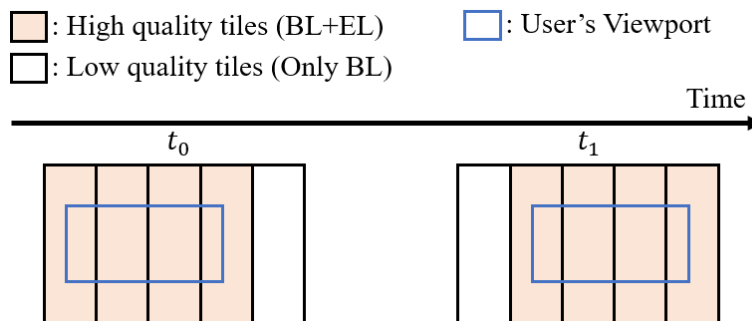
국제 비디오 코딩 표준 기술 중 스케일러블 비디오(Scalable Video)가 있다. 스케일러블 비디오 부호화 (Scalable Video Coding) 및 고효율 비디오 부호화 (High Efficiency Video Coding)의 확장 표준인 스케일러블 고효율 비디오 부호화 (Scalable High Efficiency Video Coding)는 각각 기본 계층(BL)과 하나 이상의 향상 계층 (EL)으로 이루어져 있으며, 기본 계층만 받았을 때엔 일반 화질을 제공하고 향상 계층까지 함께 받을 경우 고품질을 제공할 수 있다. 즉, 기본 계층과 하나 이상의 향상 계층이 있을 때, 기본 계층을 받은 상태에서 향상 계층을 더 받으면 받을수록 화질이나 제공하는 영상의 품질이 좋아지게 한다.

360도 영상의 높은 대역폭을 낮추기 위해, 삼성전자는 화면의 상하단 영상에 대해 다운스케일링(down-scaling)을 하여 전체 비트율(Bitrate)을 낮추고, 독일의 프라운호퍼 HHI에서는 스케일러블 고효율 비디오 부호화 (Scalable High Efficiency Video Coding)기술을 사용하되, 기본/향상 계층 모두에서 타일 기법을 사용하여 필요한 부분만 전송하고 업데이트 하는 기술을 사용한다. 하지만, 특히 후자의 기술은 빠른 사용자 응답속도를 지원하기 어렵다는 문제점을 가진다.

## 5.2. 본 표준 기술을 통해 기대되는 성능향상

### 5.2.1. 대역폭 절감효과 저하 방지

전체 영상을 하나의 압축된 영상 비트스트림(Bitstream)으로 받아서 이를 복호화(decoding)하고 사용자가 바라보는 영역을 가상의 공간에 렌더(render)하는 기술은 현재 대부분 전체 영상(예: 360도 몰입형(immersive) 영상)을 모두 비트스트림으로 전송 받는다. 각각이 고해상도인 영상이 모인 이 비디오 비트스트림의 총 대역폭은 매우 클 수밖에 없고 이를 막기 위해서 국제 비디오 표준 기술 중 SVC 및 HEVC의 스케일러블 확장 표준인 스케일러블 고효율 비디오 부호화 (Scalable High Efficiency Video Coding)와 같은 스케일러블 비디오 기술이 그림 5-2처럼 사용될 수 있다.



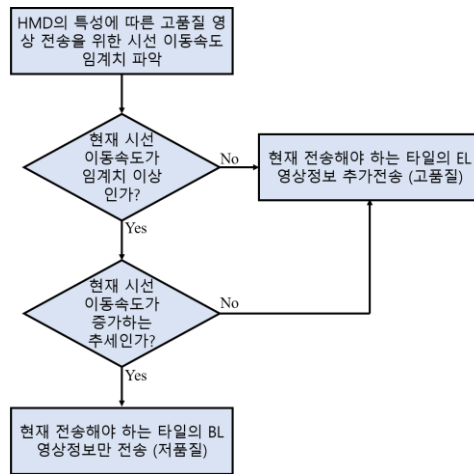
(그림 5-2) 타일과 스케일러블 비디오 기술을 활용한 사용자 시선기반 360도 영상전송

위의 그림은 사용자가 바라보고 있는 영역인 뷰포트에 해당하는 타일들은 기본품질 영상정보인 기본계층과 더불어 고품질 영상정보인 향상 계층까지 전송하고, 이외의 영역은 기본계층만 전송하는 머리장착형 영상장치를 위한 360도 영상전송 방식을 나타낸다. 이

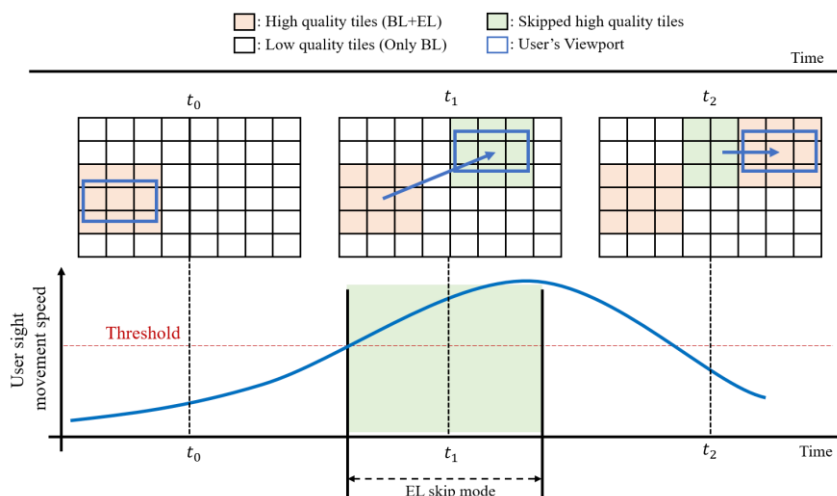


때, 사용자의 시선 이동속도는 일정하지 않으며, 사용자의 성향, 영상 콘텐츠의 특성 등에 따라 변화하게 된다. 사용자의 시선 이동속도가 느릴 때, 그림 5-2 방식은 360도 영상정보를 모두 전송하는 방식에 비하여 확실히 영상전송을 위한 대역폭을 절감할 수 있다. 하지만, 사용자의 시선변화가 잦고, 시선 이동속도가 빠르면 고품질 영상 정보를 전송하는 빈도가 높아져 스케일러블 비디오 기술 및 타일링 기법을 사용하는 의미가 약해지며 대역폭 절감효과는 저하되게 된다.

이러한 점을 해결하기 위하여 본 표준은 사용자의 시선 이동속도가 증가하게 되면 사용자의 영상품질 인지 민감도가 낮아지는 특성을 이용한다. 즉, 그림 5-3과 같이 사용자의 시선 이동속도가 특정 임계치 이상이거나 증가하고 있는 추세이면 향상계층 영상정보 전송을 생략하고(본 표준에서는 EL생략(Skip)모드라고 표기함), 사용자의 시선 이동속도가 특정 임계치 이하이거나 감소하고 있는 추세이면 향상계층 영상정보를 전송하여(본 표준에서는 EL요청(Skip)모드라고 표기함) 사용자의 빠른 시선이동으로 인한 대역폭 절감효과 저하 현상을 해결한다. 그림 5-4는 본 표준이 제안하는 사용자 시선 이동속도에 기반한 영상품질 제어방식의 예시를 나타낸 그림이다.



(그림 5-3) 사용자 시선 이동속도에 따른 전송할 영상 계층의 선택

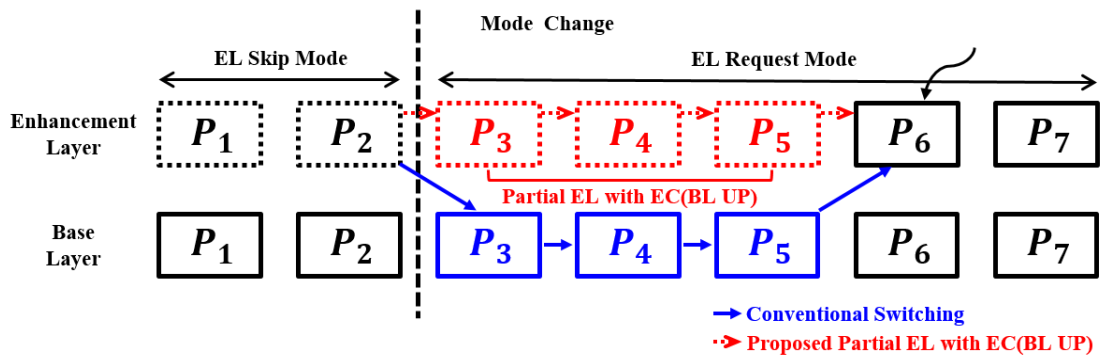


(그림 5-4) 사용자 시선 이동속도에 따른 영상품질 제어 예시

5.2.1. EL요청모드 재진입 시 에러 은닉(EC) 기술을 통한 화질 개선

기존 스케일러블 비디오 기술 및 타일링 기법이 적용된 영상전송 시스템의 경우, 현재 사용자가 바라보고 있는 뷰포트에 해당하는 타일들만 고화질 영상정보를 제공한다. 이때, 사용자의 뷰포트 이동으로 인하여 새롭게 고화질 영상정보를 제공해야 하는 타일들은 항상계층의 영상정보를 전달받아 영상 복호화를 진행해야 하는데, 영상 움직임 예측 구조의 제약 때문에 항상계층 영상들 간의 움직임 보상 및 복호화 진행이 불가능한 경우가 발생한다. 이러한 현상은 영상 내 예측 기법만 적용하여 인접 영상들과 복호화 진행에 있어 종속성이 제거된 인트라 픽처(Intra Picture, I-Picture)가 전송될 때까지 계속되며, 고화질 영상정보를 제공하기까지의 지연을 발생시킨다. 해당 지연동안 사용자는 저품질 영상정보만을 제공받을 수 있으며, 이로 인하여 멀미 등과 같은 피로감을 느낄 수 있다.

이러한 점을 해결하기 위하여, 본 표준은 그림 5-5와 같이 EL요청모드 재진입 시 참조영상(Reference Picture)의 업 샘플링 된 기본계층(Upsampled BL)을 활용하여 에러 은닉하는 기법을 사용한다. 즉, 전송되지 않아 활용할 수 없는 참조영상의 항상계층 영상정보 대신 참조영상의 기본계층 영상정보를 업 샘플링 하여 움직임 보상에 활용한다. 위에서 설명한 것처럼 기존 방식은 인트라 픽처가 전송될 때까지 고품질 영상제공에 지연이 존재하지만, 해당 기법은 해당 지연시간 동안 참조영상의 기본계층 및 현재 타일의 항상계층 영상정보를 활용하여 기본계층 영상정보만을 제공하는 경우보다 더 향상된 품질의 영상정보를 사용자에게 제공한다. 이는 평균적인 서비스 영상품질 향상 및 급격한 영상품질 변화로 인한 사용자의 멀미 등과 같은 피로감을 완화시켜주는 효과를 가져올 수 있다.



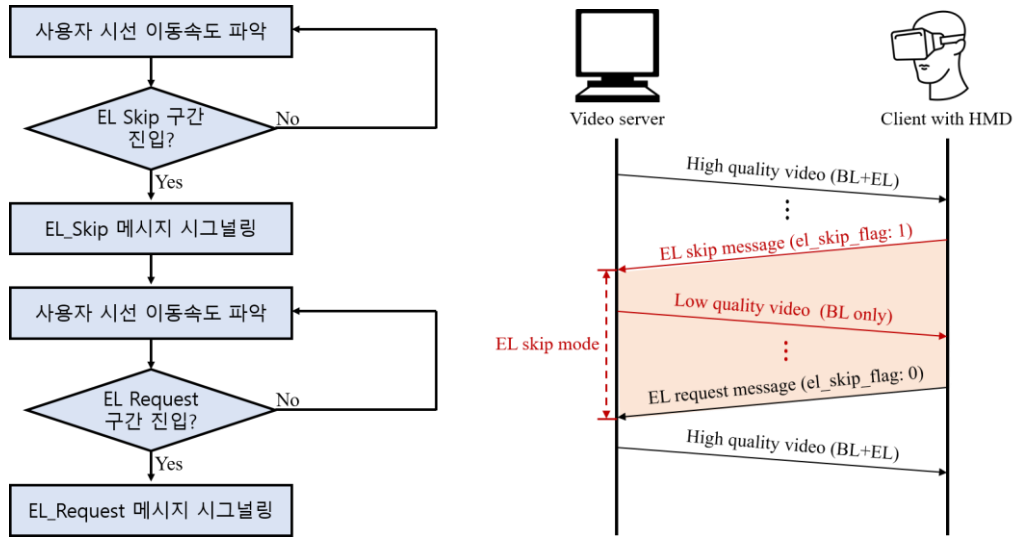
(그림 5-5) EL요청모드 재진입 시 업 샘플링 된 기본계층을 활용한 에러 은닉

### 5.3. 제안하는 표준 시그널링 규격

#### 5.3.1. 영상품질 조절정보 시그널링

앞 절에서 설명된 이유로 사용자의 시선 이동속도 기반 영상품질 조절정보를 시그널링 하는 부분은 최소한의 전송 대역폭에서 최대한의 고품질 영상 서비스를 사용자에게 제공하기 위하여 매우 중요하다. 하지만, 사용자가 시선을 이동할 때마다 그 이동속도 정보를 서버 측으로 전달하는 것은 전체 네트워크 및 사용자 단말 또는 영상 전송 시스템에 부담을 준다. 따라서, 본 발명은 그림 5-6과 같이 사용자의 시선 이

동속도가 빨라 고품질 영상정보를 생략하는 구간(본 표준에서는 EL 생략(Skip) 구간이라고 표기함)이 시작되는 시점과 사용자의 시선 이동속도가 감소하여 고품질 영상정보를 다시 요청하는 구간(본 표준에서는 EL 요청(Request) 구간이라고 표기함)이 시작되는 시점에서만 영상품질 조절정보 시그널링을 사용한다.



(a) 제안하는 시그널링 기법 (b) EL skip 및 EL request 메시지 시그널링

(그림 5-6) 본 표준의 영상품질 제어정보 시그널링

위의 시그널링을 통해 사용자의 영상품질 인지 민감도가 높아 고품질 영상정보를 인지할 수 있을 때에만 BL과 EL의 영상정보를 전송함으로써 전체적인 대역폭을 낮추고, 비디오 프로세싱 시간을 줄일 수 있다. 이는 빠른 HMD 사용자 응답시간을 보장하는데 매우 중요하다.

표 5-1은 H.264 AVC나 H.265 HEVC와 같은 국제 비디오 표준에서의 OMAF 구문(Syntax)의 예를 보여주고 있다. 표 5-1의 검은색 글씨는 기존의 표준 구문이며, 붉은색 글씨는 새로 추가되어야 할 내용이다. 표 5-2는 이의 구문 설명을 한다.

<표 5-1> 제안하는 OMAF 구문의 예

구문
aligned(8) RegionOnSphereSphereRegionStruct(range_included_flag) {
signed int(32) center_yaw;
signed int(32) center_pitch;
singed int(32) center_roll;
if (range_included_flag) {
unsigned int(32) hor_range;
unsigned int(32) ver_range;

```

}
unsigned int(1) interpolate;
bit(7) reserved = 0;
bit(1) el_skip_flag;
}
aligned(8) RegionOnSphereSphereRegionSample() {
    for (i = 0; i < num_regions; i++)
        RegionOnSphereSphereRegionStruct(dynamic_range_flag)
}
    
```

표에 나온 unsigned (n)는 통상 프로그래밍 언어에서 부호가 없는 (unsigned) ‘n’ 비트 수를 한다.

<표 5-2> 표 5-1의 구문에 대한 의미론 (Semantics)

구문	의미론
center_yaw	전역 좌표축을 기준으로 뷰포트 방향을 지정하며 뷰포트의 중심을 나타낸다. 범위는 $-180 * 2^{16} \sim 180 * 2^{16} - 1$ 내에 있어야 한다.
center_pitch	전역 좌표축을 기준으로 뷰포트 방향을 지정하며 뷰포트의 중심을 나타낸다. 범위는 $-90 * 2^{16} \sim 90 * 2^{16} - 1$ 내에 있어야 한다.
center_roll	전역 좌표축을 기준으로 뷰포트 방향을 지정하며 뷰포트의 roll 좌표를 나타낸다. 범위는 $-180 * 2^{16} \sim 180 * 2^{16} - 1$ 내에 있어야 한다.
hor_range	구 영역에서 수평 범위를 나타낸다. 구체 영역의 중심점을 통해 범위를 지정하며 $0 \sim 720 * 2^{16}$ 내에 있어야 한다.
ver_range	구 영역에서 수직 범위를 나타낸다. 구체 영역의 중심점을 통해 범위를 지정하며 $0 \sim 180 * 2^{16}$ 내에 있어야 한다.
interpolate	선형 보간의 적용 여부를 나타낸다. 값이 1일 경우 선형 보간이 적용됨을 나타낸다..
el_skip_flag	각 타일의 대응하는 영역의 EL 전송 여부를 나타내는 플래그 0: EL요청모드 (EL request mode) 1: EL생략모드 (EL skip mode)

이상 정의된 구문과 의미론에 관한 정보들은 MPEG DASH와 같은 HTTP 기반의 영상 통신에서 각각 XML 형태로 표현이 될 수도 있다. 다음 표 5-5는 XML 형태로 (1) 정보 모드, (2) 전체 타일 개수, (3) 각 타일 별 EL 전송여부 정보를 표현한 한 예이다.

&lt;표 5-5&gt; XML 형태로 표현된 타일 정보 구문

```

< tile_info>
<center_yaw= "134" center_pitch ="85" center_roll ="247" interpolate = "0" el_skip_flag = "1">
</ tile_info>

```

#### 5.4. 추가 확장적 사용

본 표준은 스케일러블 비디오를 통한 영상품질 제어기법을 이야기하고 있지만, 단일 계층을 갖는 비디오 코딩 기술을 사용할 경우에도 양자화 계수 (Quantization Parameter) 나 다운/업 스케일링 정도를 조절함으로써 전체 대역폭을 낮출 수 있는 장점을 보일 수 있다. 또한 절약된 대역폭을 활용하여 사용자의 시선 이동속도 및 경로를 파악 후, 다음에 전송할 영상정보를 미리 예측 및 전송하여 사용자의 빠른 시선이동 시에도 끊김 없는 영상전송 및 빠른 응답속도를 제공할 수 있다.

## 부 록 1-1

해당사항 없음

### 지식재산권 협약서 정보

#### 1-1.1 지식재산권 협약서

- 발명의 명칭:
- 권리자의 성명:
- 등록(출원) 번호:
- 등록(출원) 연월일:
- 실시조건:
- 협약서 접수일:

## 부 록 1-2

(본 부록은 표준을 보충하기 위한 내용으로 표준의 일부는 아님)

### 시험인증 관련 사항

해당 사항 없음

## 부 록 1-3

(본 부록은 표준을 보충하기 위한 내용으로 표준의 일부는 아님)

### 본 표준의 연계(family) 표준

해당 사항 없음



## 부 록 1-4

(본 부록은 표준을 보충하기 위한 내용으로 표준의 일부는 아님)

### 참고 문헌

해당 사항 없음

## 부 록 1-5

(본 부록은 표준을 보충하기 위한 내용으로 표준의 일부는 아님)

### 영문표준 해설서

해당 사항 없음

## 부 록 1-6

(본 부록은 표준을 보충하기 위한 내용으로 표준의 일부는 아님)

### 표준의 이력

판수	채택일	표준번호	내용	담당 위원회
제1판	2017.xx.xx	이후 기재사항	이후 기재사항	이후 기재사항