

TTA Standard

(기술보고서는 TTA Technical Report)

정보통신단체표준(국문표준)
TTAx.xx-xx.xxxx/R1

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

머리장착형 영상장치를 이용한 클라우드 기반 게임 스트리밍에서의 예상되는 사용자 뷰포트 정보 시그널링

Expected User Viewport Information
Signaling for Cloud-based Game Video
Streaming using Head-Mounted Display
(HMD)

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	노현준	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 확약서 정보는 본 표준의 '부록(지식재산권 확약서 정보)'에 명시하고 있으며, 이후 접수된 지식재산권 확약서는 TTA 웹사이트에서 확인할 수 있습니다.

본 표준과 관련하여 접수된 확약서 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) 및 클라우드 기반 게임 스트리밍(Cloud Based Game Streaming) 기술과 머리장착형 영상장치(Head-Mounted Display; HMD)를 위한 게임콘텐츠 전송시 일부 우선시되는 영역의 정보를 시그널링(Signaling) 함으로써, 비디오 요구 대역폭을 낮추고, 주관적 화질을 유지하면서 저지연을 달성하는 기술을 설명함에 있다.

2 주요 내용 요약

이 표준은 머리장착형 영상장치를 통한 게임 스트리밍 콘텐츠를 전송할 때, 사용자의 시선이 향할 확률이 높은 영역의 정보를 시그널링하는 기술을 담고 있으며, 기술의 장점을 설명하기 위해 (1) 본 표준의 필요성과 스케일러블 비디오 및 머리장착형 영상장치의 배경지식, (2) 본 표준을 머리장착형 영상장치를 활용한 게임 스트리밍 콘텐츠 전송에 적용했을 때 기대되는 성능향상, 그리고, (3) 제안하는 표준 시그널링 규격(신텍스, 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 syntax and semantics of the user sight information for VR360 cloud-based game video streaming using HMD. The signaling standard can decrease the video adaptation delay for game video streaming as well as the bandwidth requirements of video transmission.

2 Summary

The standard includes the specifications for the user sight 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, 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 이용한 클라우드 기반 게임 스트리밍에서의 예상되는 사용자 뷰포트 정보 시그널링	2
5.1 본 표준의 필요성	2
5.2 본 표준 기술을 통해 기대되는 성능향상	3
5.3 제안하는 표준 시그널링 규격	4
5.4 추가 확장적 사용	9
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 이용한 클라우드 기반 게임 스트리밍에서의 예상되는 사용자 뷰포트 정보 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) 비디오 표준의 SEI, VUI, 또는 슬라이스 헤더 (Slice Header) 등의 패킷 단위에서 전해질 수도 있고, (3) 비디오 파일을 설명(Describe)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
- HEVC: High Efficiency Video Coding
- SVC: Scalable 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

5 머리장착형 영상장치를 이용한 클라우드 기반 게임 스트리밍에서의 예상되는 사용자 뷰포트 정보 시그널링

본 섹션에서는 머리장착형 영상장치(HMD)를 이용한 게임 스트리밍에서의 예상되는 사용자 뷰포트(Viewport) 정보 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 대표적인 영화 관람 및 게임뿐만 아니라 화상회의와 원격 수술 등이 존재한다. 이 중 게임 같은 콘텐츠들은 일반 사용자들이 쉽게 접할 수 있고 머리장착형 영상장치를 구입하는 요인이 되는 콘텐츠이다.

클라우드 기반 게임 스트리밍 역시 널리 보급되고 있는데, 이는 서버에서 게임과 관련된 주요 연산들이 처리되고 클라이언트는 서버에 접속하여 게임 화면을 전송 받아 게임을 즐기는 기술이다. 이 기술은 클라이언트의 연산 성능에 제약 없이 고사양 게임을 즐기는 장점이 존재한다.

아직까지 머리장착형 영상장치와 클라우드 기반 게임 스트리밍을 함께 사용하는 것이 불편적이지는 않지만, 몰입도가 상당한 머리장착형 영상장치의 특징과 앞서 언급한 클라우드 게임 스트리밍의 장점으로 인해 근시일 내에 널리 보급될 것이라 예상된다.

현재, 머리장착형 영상장치를 이용한 콘텐츠들이 겪는 어려운 문제는 사용자 눈에 매우 넓게 보이는 (1) 360 영상 전체를 담는 비디오 화소수가 매우 높아야 한다는 점이다. 따라서 UHD 급 영상을 이용할 필요성이 있는데, 이 경우 사용자 단말들 사이에 (2) 대역폭 확보가 어려운 문제점과, 처리해야 할 많은 비디오 데이터로 인해 (3) 사용자의 머리 움직임에 빠르게 응답하기 어렵다는 문제점을 갖게 된다. 특히, 클라우드 게임 콘텐츠들은 콘텐츠 특성상 인/디코딩 과정이나 전송과정에서 지연이 발생하므로 필요 대역폭 축소와 즉각적인 반응이 더욱 중요하다.

본 표준은 이의 해결책으로 ‘예상되는 사용자 뷰포트 정보 시그널링’ 기법의 내용과 장점, 그리고 이를 통한 (1) 비디오 프로세싱 속도 향상 및 전체 대역폭 낮춤, (2) 대역폭에 따른 전송 우선순위 지정을 설명한다.

5.1.1. 배경지식

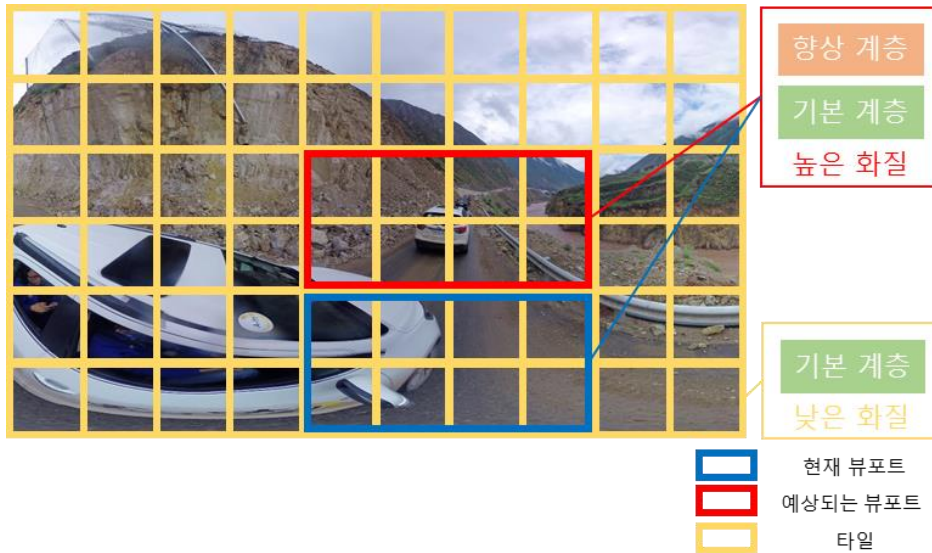
국제 비디오 코딩 표준 기술 중 스케일러블 비디오(Scalable Video)가 있다. 스케일러블 비디오 부호화(Scalable Video Coding) 및 고효율 비디오 부호화(High Efficiency Video Coding)의 확장 표준인 스케일러블 고효율 비디오 부호화(Scalable High Efficiency Video Coding)는 각각 기본 계층(BL)과 하나 이상의 향상 계층(EL)으로 이루어져 있으며, 기본 계층만 받았을 때엔 일반 화질을 제공하고 향상 계층까지 함께 받을 경우 고품질을 제공할 수 있다. 즉, 기본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

받은 상태에서 항상 계층을 더 받으면 받을수록 화질이나 제공하는 영상의 품질이 좋아지게 한다.

5.2. 본 표준 기술을 통해 기대되는 성능향상

5.2.1. 비디오 프로세싱 속도와 대역폭 문제의 해결

전체 영상을 하나의 압축된 영상 비트스트림(Bitstream)으로 받아서 이를 복호화(Decoding)하고 사용자가 바라보는 영역을 가상의 공간에 렌더링(Rendering)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 사용될 수 있다. 그 방법은 그림 5-1과 같다.



(그림 5-1) 예상되는 뷰포트를 선반입하는 게임 스트리밍 방식

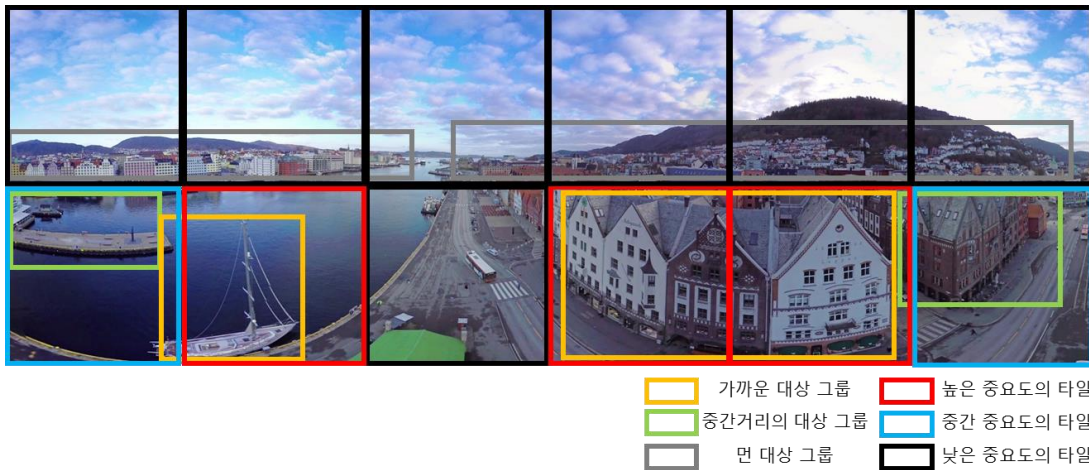
위의 그림은 사용자가 클라우드 기반 게임 스트리밍을 통해 카레이싱을 할 때를 가정한 것이다. 파란 상자는 현재 보여지고 있는 화면의 중앙부로 현재 사용자의 뷰포트이다. 이 부분은 사용자가 현재 보고 있는 영역이므로 높은 화질이 요구되고, 따라서 스케일러블 비디오 부호화의 기본 계층과 향상 계층 모두 전송해준다. 그림에서 예로 사용한 카레이싱의 경우, 예상되는 뷰포트는 정면의 주행코스로 예상할 수 있다. 따라서 그 예상되는 뷰포트를 미리 높은 화질로 선반입(Prefetching)함으로써 지연을 줄일 수 있게 한다. 또한 현재 뷰포트나 예상되는 뷰포트 이외의 구간은 기본 계층 정보만 보내줌으로써 전체 대역폭을 낮출 수 있다. 물론 기본 계층으로만 구성된 부분들의 경우도 멀미를 저감하도록 일정 수준 이상의 화질을 유지해야 한다.

5.2.2. 대역폭에 따른 전송 우선순위 지정

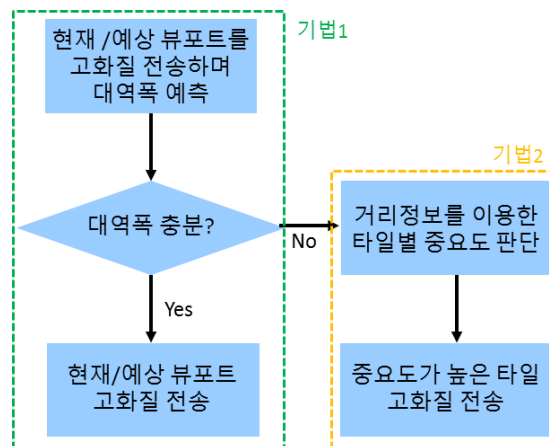
클라우드 기반 게임 스트리밍의 주요과제는 앞서 언급했듯이 저지연 스트리밍이다. 그런데 머리장착형 영상장치의 경우 무선네트워크를 이용하는 방식이기 때문에 대역폭 변동의 가능성이 항상 존재한다. 만약 특정영역을 항상 높은 화질로 전송하도록 고정해 버린다면, 갑자기 대역폭이 줄어들었을 때 지연이 발생할 수도 있다.

본 표준 기술은 게임화면 내 대상(Object)들의 거리정보를 이용하여 어떤 타일이 가까운 대상들을 많이 포함하고 있는지 구분한 뒤, 가까운 타일(가까운 대상이 많은 타일) 순으로 정렬하도록 한다. 이후에 대역폭이 갑자기 감소하게 되면, 대역폭 허용 범위까지 가까운 타일 순으로 고화질 전송하여 지연 없이 스트리밍이 전송될 수 있도록 한다. 따라서 대역폭이 충분하다면 전부 고화질로 전송이 가능할 것이고, 대역폭이 충분하지 않다면 가까운 대상들만 고화질로 전송이 이뤄질 것이다. 거리가 가까운 대상일수록 뷰포트 가능성이 높기 때문에 이러한 기법을 도입한다. 고화질/저화질 영상을 만드는 방식은 5.2.1과 마찬가지로 스케일러블 비디오 기술을 적용한다. 해당 기술은 주인공의 주변 사물들이 우선시되는 역할 게임(Role-Playing Game; RPG)에서 더욱 효과적일 것으로 예상된다.

앞서 소개한 5.2.1과 5.2.2의 기법들을 각각 ‘기법1’, ‘기법2’라고 표현하면 그 적용 순서도는 아래 그림 5-3과 같다.



(그림 5-2) 거리에 따라 뷰포트 예측 가능성을 구분하여 고화질로 전송하는 게임 스트리밍 방식



(그림 5-4) 본 표준의 기법 적용 순서도

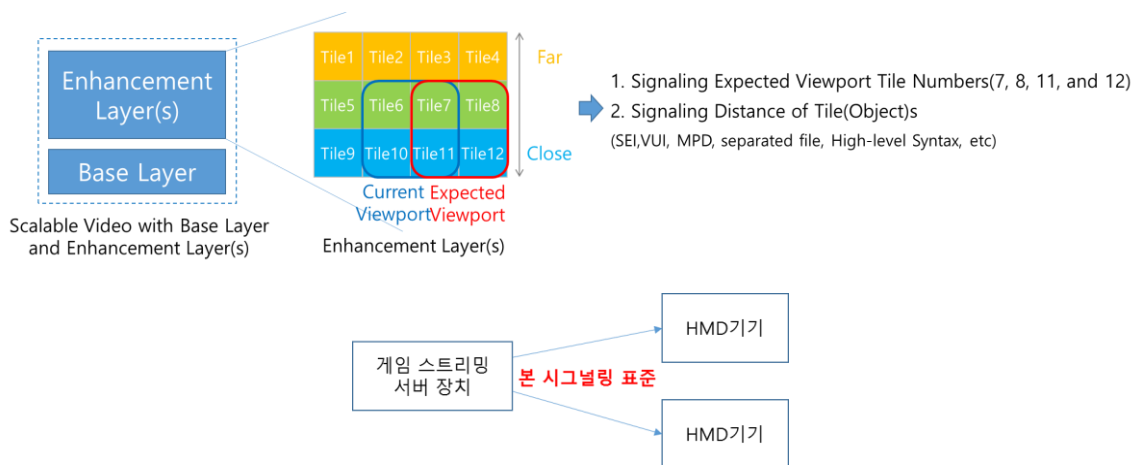
5.3. 제안하는 표준 시그널링 규격

5.3.1. 예상되는 뷰포트(Viewport) 및 각 대상(Object)들의 거리 정보 시그널링

게임 제작자는 사용자의 시선이 ‘마땅히 주목 받을 것으로 예상되는 (본 표준에서는 Viewport로 표기함)’ 오브젝트를 미리 알 수 있도록 시그널링 하여, 사용자의 시선이 향하기 전에 미리 항상 계층까지 선반입 할 수 있도록 한다. 또한 각 타일들의 거리 정보도 옵션으로 전송되어 상황에 따라 사용될 수 있도록 한다. 각 대상에 따라 거리 정보를 전송하는 것이 가장 이상적이지만, 전송하는 화면의 최소 구성 단위가 대상 하나하나가 아닌 타일이기에 각 대상들의 거리정보는 각 타일에 통합되어져서 처리된다.

기본 계층은 빠른 사용자 응답시간을 위해 타일링되지 않고 전체적으로 인코딩된다. 하나 이상의 향상계층은 필요에 따라 일부 또는 전체가 여러 타일들로 나누어져서 인코딩된다. 이 때, 뷰포트는 (1) 사용자의 시선이 바라보는 곳, 또는 (2) 사용자가 보게 될 가상의 공간에서 중요 오브젝트가 위치할 타일 위치 (예: 게임 등에서 새로운 적이 등장하는 위치에 해당할 수 있다).

뷰포트의 타일 번호는 연속적이므로 모든 번호를 다 보내지 않고도 효과적으로 압축할 수 있다 (예: 타일 시작과 끝 번호, 좌표점 정보, 타일 내 코딩 단위(CU) 번호 리스트, 타일 번호를 수식으로 표현 등).



(그림 5-4) 본 표준의 예상되는 뷰포트와 대상 거리 정보 시그널링 예

이 시그널링은 그림 5-4와 같이 (1) 세션(Session) 정보를 실어나르는 고수준 구문 (High-level syntax) 프로토콜을 통해 전해질 수도 있고, (2) 비디오 표준의 SEI, VUI, 또는 슬라이스 헤더 등의 패킷 단위에서 전해질 수도 있고, (3) 비디오 파일을 설명하는

별도의 파일로(예: DASH의 MPD) 전달될 수 있다.

위의 시그널링을 통해 향상계층의 특정 타일만 우선적으로 전달 받아 전체적인 지연 시간을 줄이고, 대역폭 상황에 따라 일부만 고화질로 처리함으로써 지연발생을 없앨 수 있다. 이는 빠른 HMD 사용자 응답시간을 보장하는데 매우 중요하다.

표 5-1은 H.264 AVC나 H.265 HEVC와 같은 국제 비디오 표준에서의 SEI 메시지 페이로드 (payload) 구문 (Syntax)의 예 (expected_tile_info)를 보여주고 있다. 만일 제안하는 구문이 188번으로 정해진 경우 다음과 같다. 검은색 글씨는 기존의 표준 구문이며, 붉은색 글씨는 새로 추가되어야 할 내용이다.

<표 5-1> 제안하는 SEI 페이로드 구문의 예

Sei_payload(payloadType, payloadSize) {	Descriptor
if(payloadType == 0)	
buffering_period(payloadSize)	
.....	
else if(payloadType == 188)	
expected_tile_info(payloadSize)	
.....	

다음 표 5-2와 표 5-3은 본 표준에서 제안하는 비디오 픽처별 뷰포트(시선 정보)(Viewport) 시그널링 규격과 파일, 청크, 비디오 픽처 그룹별 뷰포트(시선 정보)(Viewport) 시그널링 규격을 다루고 있고, 표 5-4는 이의 구문 설명을 한다.

표에 나온 u(n)는 통상 프로그래밍 언어에서 부호가 없는 (unsigned) ‘n’ 비트 수를 의미하며, ‘v’로 표시된 부분은 변화 가능한 비트수(표준에서는 varies로 읽힘)를 의미한다.

<표 5-2> 제안하는 비디오 픽처별 시그널링 규격

expected_tile_info (payloadSize) {	비트수
info_mode	u(4)
viewport_num	u(8)
tile_num	u(12)
if (info_mode == 1) {	
for (i=0; i < viewport_num; i++) {	
tile_id_list_in_viewport[i]	u(12)
}	
}	
if (info_mode == 2) {	

viewport_id_list_trans[viewport_num]	u(12)
tile_id_list_trans[unit_num]	u(12)
}	
if (info_mode == 3) {	
for (i=0; i < viewport_num; i++) {	
tile_distance_list_in_viewport[i]	u(16)
}	
}	
user_info_flag	u(1)
if (user_info_flag) {	
user_info_size	u(8)
for (i=0; i < user_info_size; i++) {	
user_info_list [i]	u(v)
}	
}	
}	

<표 5-3> 제안하는 파일, 청크, 비디오 픽처 그룹별 시그널링 규격

expected_tile_info_file {	비트수
version_info	u(8)
file_size	u(64)
for (j=0; j < file_size; j++) {	
poc_num	u(32)
info_mode	u(4)
viewport_num	u(8)
tile_num	u(12)
if (info_mode == 1) {	
for (i=0; i < viewport_num; i++) {	
tile_id_list_in_viewport[i]	u(12)
}	
}	
if (info_mode == 2) {	
viewport_id_list_trans[viewport_num]	u(12)
tile_id_list_trans[unit_num]	u(12)
}	
if (info_mode == 3) {	

for (i=0; i < viewport_num; i++) {	
tile_distance_list_in_viewport[i]	u(16)
}	
}	
user_info_flag	u(1)
if (user_info_flag) {	
user_info_size	u(8)
for (i=0; i < user_info_size; i++) {	
user_info_list [i]	u(v)
}	
}	
}	
}	

<표 5-4> 표 5-2와 5-3의 구문에 대한 의미론 (Semantics)

구문	의미론
version_info	시그널링 규약의 버전 정보, 부호 없는 8비트의 정보로 표현된다.
file_size	파일 사이즈, 부호 없는 64 비트의 정보로 표현된다.
poc_num	HEVC와 같은 비디오 표준에서의 POC (Picture Order Count) 정보를 의미함, 기존의 H.264 AVC 표준에서의 프레임 번호 (frame number)와 유사한 의미. 부호 없는 32 비트의 정보로 표현된다.
info_mode	본 표준에서 정의한 ‘정보 모드’로서 다음과 같으며, 부호 없는 4 비트의 정보로 표현된다. 0: 이전 시그널링 정보와 같음 1: 예측되는 각 뷰포트에 포함되는 타일 id 2: 예측되는 각 뷰포트에 포함되는 타일에 대한 거리정보 3: 전송되는 뷰포트 id 및 타일 id
viewport_num	예측되는 뷰포트 갯수, 부호 없는 8 비트의 정보로 표현된다.
tile_num	화면 내의 타일 갯수, 부호 없는 12 비트의 정보로 표현된다.
tile_id_list_in_viewport[]	뷰포트 내 타일 번호 리스트 , 부호 없는 12 비트의 정보로 표현된다.
tile_distance_list_in_viewport[]	뷰포트 내 타일 별 거리 정보 리스트, 각각의 거리 정보는 부호 없는 16 비트의 정보로 표현된다.

viewport_id_list_trans[]	전송되는 뷰포트 번호 리스트 , 부호 없는 12 비트의 정보로 표현된다.
tile_id_list_trans[]	전송되는 타일 번호 리스트, 부호 없는 12 비트의 정보로 표현된다.
user_info_flag	추가 사용자 정보 모드의 플래그 (flag), 사용자가 추가로 전송하려는 타일 관련 정보가 있는지 여부가 부호 없는 1 비트의 정보로 표현된다. 0: 추가 사용자 정보가 없음. 1: 추가 사용자 정보가 있음.
user_info_size	추가 사용자 정보의 길이, 부호 없는 16 비트의 정보로 표현한다.
user_info_list []	추가 사용자 정보의 리스트, 각각의 추가 사용자 정보는 부호 없는 변화 가능한 (varies) 비트의 정보로 표현된다.

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) 추가 사용자 정의 모드 플래그, (3) 뷰포트 갯수 정보, (4) 타일 갯수 정보, (5) 전송되는 뷰포트 번호 정보, (6) 전송되는 타일 번호 정보를 표현한 한 예이다.

<표 5-5> XML 형태로 표현된 타일 정보 구문

```
< expected_tile_info>
<info_mode = "3 user_info_flag ="0" viewport_num ="2" tile_num "6" viewport_id_list_trans ="1 2"
tile_id_list_trans ="4 5 12 14 17 22">
</ expected_tile_info>
```

5.4. 추가 확장적 사용

본 표준은 스케일러블 비디오와 뷰포트 및 거리정보를 통한 차별적 전송 기법을 이야기 하고 있지만, 화면 분할을 지원하는 다른 비디오 병렬처리 기법들(예: 슬라이스 (Slice), FMO (Flexible Macro Block) 등)에 적용 가능하다. 또한 비트 스트림을 분할하여 전송하는 스트리밍 서비스인 MPEG DASH, 마이크로소프트(MS)사의 Smooth 스트리밍 (Streaming), 애플(Apple)사의 HLS (HTTP Live Streaming; HTTP 라이브 스트리밍)에 적

용 가능하다.

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	2016.xx.xx	이후 기재사항	이후 기재사항	이후 기재사항