

# TTA Standard

정보통신단체표준(국문표준)

TTAK.KO-10.1392

제정일: 2022년 12월 07일

적응적 몰입형 비디오 전송을 위한  
확장 부호화 부가정보 구성 요소 및  
형식

Syntax and Semantics of Scalable Coding Metadata for  
Adaptive Transmission of Immersive Video



한국정보통신기술협회  
Telecommunications Technology Association

**표준초안 검토 위원회 디지털콘텐츠 프로젝트그룹(PG610)**

**표준안 심의 위원회 소프트웨어/콘텐츠 기술위원회(TC6)**

	성명	소 속	직위	위원회 및 직위
<b>표준(과제) 제안</b>	류은석	성균관대학교	부교수	PG610 특별위원
	이순빈	성균관대학교	대학원생	-
	정종범	성균관대학교	대학원생	-
다차원영상기술표준화포럼				
<b>표준 초안 에디터</b>	류은석	성균관대학교	부교수	PG610 특별위원
	김상원	한국전자통신연구원	책임연구원	PG610 간사
	이범렬	한국전자통신연구원	책임연구원	PG610 부의장
	김영민	한국전자기술연구원	책임연구원	PG610 부의장
	남현우	동덕여자대학교	교수	PG610 의장
<b>사무국 담당</b>	김찬영	TTA	선임연구원	-

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. 준용표준인 경우 해당 표준화기구 또는 단체의 웹사이트에서 이를 확인해야 합니다.

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 : 2022. 12. 07.

# 서 문

## 1 표준의 목적

이 표준의 목적은 몰입형 영상을 위한 가상현실(virtual reality, VR) 기술을 응용하여 확장 부호화 표준 기술을 통한 비트스트림 생성 및 적응적 전송에 필요한 메타데이터를 정의함으로써 가상현실 영상 전송 시 단말과 서비스 환경에 따른 적응적인 해상도와 품질을 가진 몰입형 영상 렌더링을 제공하는 데 있다.

## 2 주요 내용 요약

이 표준은 사용자의 머리 움직임 추적이 가능한 머리 장착형 영상장치(head-mounted display, HMD)를 통한 초다시점 가상현실 영상 전송 및 렌더링 시, 확장 부호화 표준에서의 기본계층(Base Layer)과 향상계층(Enhancement Layer)간의 효율적 전송을 위한 메타데이터 기록 기술, 표준 신호 체계 규격(구문과 의미론)을 기술한다. 본 표준은 확장 부호화 기반 다시점 가상현실 영상 전송 시스템에 적용이 가능하다. 이 표준은 국제 표준단체 MPEG-Immersive 표준 기술 등을 이용하는 시스템을 위한 별도의 독립적인 시그널링 표준으로서, MPEG 비디오 시스템 표준과 직접적인 관련성이 없다.

## 3 인용 표준과의 비교

해당사항 없음.

## Preface

### 1 Purpose

The purpose of this standard is to describe technologies that enable scalable immersive video transmission with adaptive resolution and quality depending on the service environment, while organizing the required information into metadata considering adaptive transmission the bandwidth and device.

### 2 Summary

This standard describes the technical and standard specifications (syntax and semantics) about the scalable video coding with immersive video and adaptive streaming in VR video tile based transmission and rendering through a head-mounted display (HMD) devices capable of processing base layer and enhancement layer. This standard can be applied to the scalable VR video streaming system. This standard does not directly affect to or influenced by the referenced MPEG-Immersive standard but specifies the signaling details independently.

### 3 Relationship to Reference Standards

None.

## 목 차

1	적용 범위	1
2	인용 표준	1
3	용어 정의	1
4	약어	1
5	적응적 몰입형 비디오 전송을 위한 확장 부호화 부가정보 구성 요소 및 형식	2
5.1	확장 부호화 몰입형 비디오 부가 정보 구성 요소 및 형식	2
5.2	신호 체계 규격	4
부록 I	필요성	6
부록 II-1	지식재산권 요약서 정보	7
II-2	시험인증 관련 사항	8
II-3	본 표준의 연계(family) 표준	9
II-4	참고 문헌	10
II-5	영문표준 해설서	11
II-6	표준의 이력	12

# 적응적 몰입형 비디오 전송을 위한 확장 부호화 부가정보 구성 요소 및 형식

## (Syntax and Semantics of Metadata for Adaptive Transmission with Scalable Immersive Video)

### 1 적용 범위

본 표준의 적용 범위는 영상 전송에서의 부가 정보를 처리하는 객체를 다루며, 이는 사용자 단말, 서버, 중계 시스템 및 라우터 등을 포함한다. 또한, 본 표준의 부가 정보 구문(Syntax) 및 의미론(Semantics)은 세션(Session) 정보를 포함하는 고수준 구문(High-level Syntax, HLS) 프로토콜을 통해 전해질 수도 있고, 비디오 표준의 SEI, VUI 또는 슬라이스 헤더(Slice Header) 등의 패킷 단위에서 전해질 수도 있고, 비디오 파일을 설명(Descript)하는 별도의 파일로(예: DASH MPD) 전달될 수 있다.

### 2 인용 표준

해당 사항 없음

### 3 용어 정의

#### 3.1 아틀라스 (Atlas)

MPEG immersive video (MIV) 및 video-based point cloud coding (V-PCC) 기반 다수의 몰입형 영상 압축 시 생성되는 기본 시점 및 그 외 시점에서의 잔차 영상이 포함된 영상.

#### 3.2 Exponential Golomb Code

규칙적인 구조를 가지는 가변 길이 코드로, 영상 압축 시 엔트로피 부호화 단계에서 사용됨.

### 4 약어

DASH	Dynamic Adaptive Streaming over HTTP
HEVC	High-Efficiency Video Coding
SHVC	Scalable High efficiency Video Coding

MPD	Media Presentation Description
MPEG	Moving Picture Experts Group
SEI	Supplemental Enhancement Information
VVC	Versatile Video Coding

## 5 적응적 몰입형 비디오 전송을 위한 확장 부호화 부가정보 구성 요소 및 형식

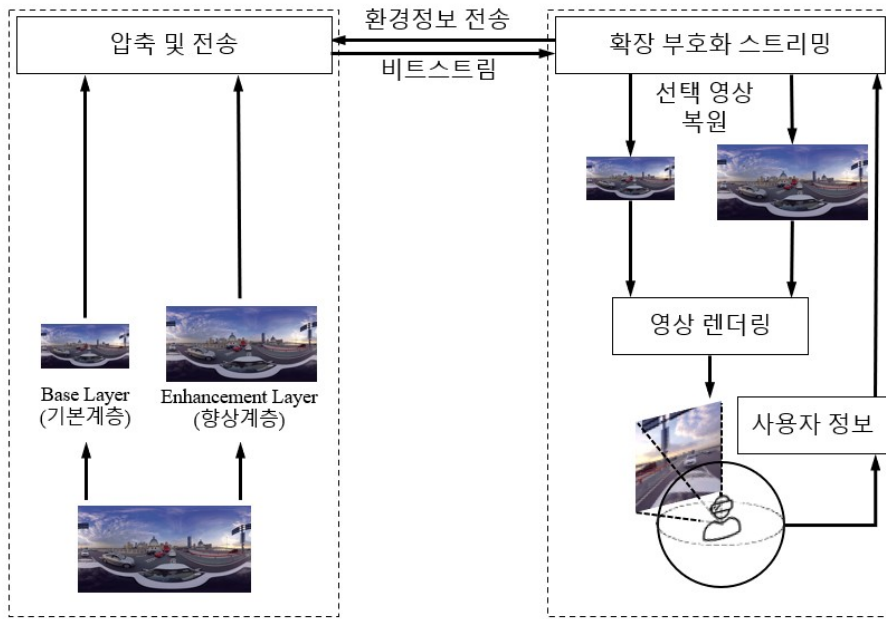
### 5.1 확장 부호화 몰입형 비디오 부가 정보 구성 요소 및 형식

#### 5.1.1 확장 부호화 몰입형 비디오 스트리밍 시스템

머리 장착형 영상장치를 통해 몰입형 가상현실을 위한 360도 영상 스트리밍 시 사용자의 단말과 서비스 환경 상태를 고려하여, 적응적인 해상도와 품질의 영상을 전송하는 기술이 사용될 수 있다. 특히, 몰입형 비디오 표준화 기술은 단말에서의 원활한 재생 및 전송을 위해 다시점 간의 중복성을 제거하는 기술 등 모바일을 우선적으로 고려한 전송 기술을 채택하고 있다. 이러한 맥락으로 몰입형 영상에서 모바일에서 원활한 몰입형 비디오 스트리밍을 가능하게 하는 적응적 해상도 및 품질 선택 정보를 포함하는 SEI 메시지를 정의하여, 몰입형 비디오에서 효율적인 SHVC 압축 및 전송이 가능하다.

처리된 몰입형 비디오의 영상에서 각 계층 별 정보를 포함하여, 확장 부호화 전송시에 해당하는 계층 별 영상과, 단말에 따라 적응적으로 향상 계층에 대한 정보를 포함하는 SEI 메시지를 정의하여, 몰입형 비디오 기술에서도 확장 부호화 압축 및 전송이 가능하다. 해당 시스템은 HEVC 및 차세대 영상 압축 도구인 VVC 표준에 맞게 확장 부호화 및 전송을 통하여 적응적인 스트리밍 환경을 제공할 수 있다.

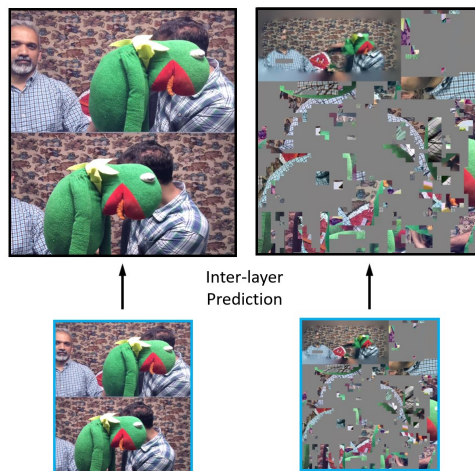
(그림 5-1)은 적응적 몰입형 비디오 전송을 위한 확장 부호화 몰입형 비디오 스트리밍 시스템 구조도를 나타낸다. 적응적인 해상도 전송을 위해 기존의 영상을 다운샘플링한 기본 계층과 향상 계층으로 분리하여 영상처리를 시행한다. 이때, HEVC의 SHVC 표준 기술 등을 통하여 계층 간 예측 기술 등이 각 계층 영상에 적용될 수 있다. 단말 또는 서비스 환경 변경 시 적응적인 전송을 위해 클라이언트에서 단말 또는 대역폭과 같은 서비스 환경에 대한 정보를 추출하여 전송한다. 중앙 서버는 전송받은 서비스 환경을 통해 기본 계층 또는 향상 계층을 포함하는 서비스 영역을 선택하고 전송한다. 이후 클라이언트는 서버로부터 전송받은 비트스트림을 복호화하고, 사용자 시점을 생성하여 머리 장착형 영상장치에 표현한다.



(그림 5-1) 적응적 몰입형 비디오 전송을 위한 확장 부호화 몰입형 비디오 스트리밍 시스템 구조도

### 5.1.2 아틀라스 수준 계층 정보 포함

몰입형 비디오는 각 다시점 영상의 중복성을 제거하고 남은 아틀라스라는 이름의 영상으로 처리를 거치게 된다. 확장 부호화 몰입형 비디오 스트리밍 시스템을 위해서는, 생성된 아틀라스 영상에 대해 각 계층에 해당하는 데이터와, 계층 간 예측을 수행하였을 때의 파라미터를 포함하는 메타데이터를 포함해야 한다. (그림 5-2)는 몰입형 비디오에서 계층 간 예측을 수행하였을 때의 예시를 나타낸다.

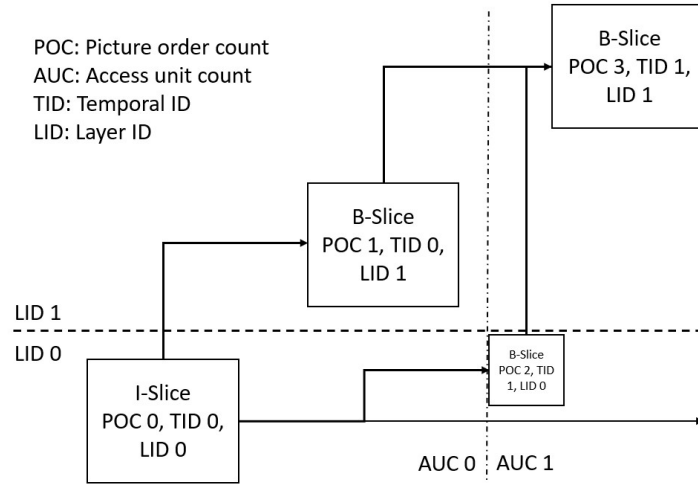


(그림 5-2) 아틀라스 별 계층 간 예측 수행 예시

몰입형 비디오는 하나의 영상이 아닌 다수의 영상으로 이루어져 있으므로 (그림 5-3)에



서 나타나듯이, 이에 대한 계층 정보를 표시하는 메타데이터를 정의해 확장 부호화 몰입형 비디오 스트리밍이 개시될 수 있다.



(그림 5-3) 확장 부호화에 대한 정보 예시

## 5.2 신호 체계 규격

본 표준의 핵심 신호 체계는 서버가 가상현실 몰입형 비디오 스트리밍 중앙 서버로부터 전달받는 비트스트림에 대한 계층 정보를 포함한다.

본 표준에서 다루는 신호 체계는 상위 수준 구문 프로토콜을 통해 전해질 수 있고, SEI 메시지를 통해 전해질 수도 있으며, 영상 파일을 설명하는 별도의 파일(예: MPEG DASH의 MPD)로 전달될 수도 있다.

다음 <표 5-1>은 몰입형 비디오 SEI 메시지에서 정의되어야 할 신호 체계 규격이다. 표에 나온  $ue(v)$ 는 통상 프로그래밍 언어에서 부호가 없는(unsigned) exponential golomb code로 부호화된 정수를 의미한다.

<표 5-1> 몰입형 비디오 SEI 메시지에서 변경되는 신호 체계 규격

신호 체계 규격	비트 수
Scalability_info {	
scalability_flag	u(1)
num_layer_minus1	ue(v)
...	
for (i=0; i<num_layer_minus1; i++)	
layer_id[i]	ue(v)
}	

신호 체계 규격	비트 수
...	
}	

다음 <표 5-2>는 몰입형 비디오 SEI 메시지에서 수정되어야 할 신호 체계에 대한 구문 의미론이다.

<표 5-2> 몰입형 비디오 SEI 메시지에서 변경되는 정보에 대한 구문 의미론

구문	의미론
scalability_flag	몰입형 비디오에서 현재 확장 부호화를 사용하고 있는지에 대한 플래그를 의미함.
num_layer_minus1	확장 부호화시에 계층 수를 의미함. 부호 없는 정수로 표현됨.
layer_id	각 영상에 해당하는 계층 정보를 의미함. 부호 없는 정수로 표현됨.

## 부 록 I

(본 부록은 표준을 보충하기 위한 내용으로 표준의 일부는 아님)

### 필요성

#### 1.1 본 표준의 필요성

최근 가상현실에 대한 관심이 증대되면서, 가상현실 관련 기술이 활발히 개발되고 있다. ISO/IEC 산하 국제 표준화 단체인 moving picture experts group(MPEG)에서 논의 중인 MPEG Immersive Video (MIV) 표준화 기술에서는 초다시점 가상현실 영상에 대한 기술들이 발전하고 있다. 특히 기존에 연구되었던 다시점 영상과는 다르게 모바일 등의 단말에서의 감상 시나리오를 고려하여 디코더 오버헤드와 같은 필요한 복잡도를 검토하면서 표준화를 진행하고 있다.

본 표준에서는 MIV의 발전 방향과 더불어 다양한 미디어 환경에서 적응적인 해상도로 영상을 전송할 수 있는 확장 부호화 기술을 몰입형 영상에 통합하고자 하였다. 몰입형 영상에서 생성되는 영상은 하나가 아닌 다수의 영상을 포함하고 있으며, 이로 인해 스트리밍 시스템에서 각 확장 부호화 기술의 계층 정보를 포함해야 할 필요가 있다. 본 표준에서 제시하는 신호 체계 규격을 통해 계층 간 예측 기술 역시 몰입형 비디오 스트리밍에서 효율적으로 사용할 수 있다. 본 표준에서 제시하는 계층 별 정보를 포함하여 몰입형 영상에 대해 확장 부호화 기술을 적용할 수 있으며, 모바일 등 다양한 단말에서 적응적으로 원활한 감상을 가능하게 한다. 따라서, 본 표준은 몰입형 비디오 영상처리를 통해 생성된 영상들을 확장 부호화로 제어하기 위한 방법 및 신호 체계 규격을 정의 및 기술한다.

## 부 록 II-1

(본 부록은 표준을 보충하기 위한 내용으로 표준의 일부는 아님)

### 지식재산권 확약서 정보

아래에 기재된 지식재산권 확약서 이외에도 본 표준이 발간된 후 접수된 확약서가 있을 수 있으니, TTA 웹사이트에서 확인하시기 바랍니다.

해당 사항 없음.

## 부 록 II-2

(본 부록은 표준을 보충하기 위한 내용으로 표준의 일부는 아님)

### 시험인증 관련 사항

해당 사항 없음.

## 부 록 II-3

(본 부록은 표준을 보충하기 위한 내용으로 표준의 일부는 아님)

### 본 표준의 연계(family) 표준

II-3.1 TTAK.KO-10.1198, 3DoF+ 영상 처리 및 전송을 위한 부가정보 구성 요소 및 형식

## 부 록 II-4

(본 부록은 표준을 보충하기 위한 내용으로 표준의 일부는 아님)

### 참고 문헌

아래 기재된 참고 문헌의 발간일이 기재된 경우, 해당 표준(문서)의 해당 버전에 대해서만 유효하며, 연도를 표시하지 않은 경우에는 해당 표준(권고)의 최신 버전을 따른다.

해당 사항 없음.

## 부 록 II-5

(본 부록은 표준을 보충하기 위한 내용으로 표준의 일부는 아님)

### 영문표준 해설서

해당 사항 없음.



## 부 록 II-6

(본 부록은 표준을 보충하기 위한 내용으로 표준의 일부는 아님)

### 표준의 이력

판수	채택일	표준번호	내용	담당 위원회
제1판	2022.12.07	제정 TTAK.KO-10.1392	-	디지털콘텐츠 프로젝트그룹 (PG610)