

TTA Standard

정보통신단체표준(국문표준)

TTAK.KO-10.1258

제정일: 2020년 12월 10일

가상현실 영상 다중 타일 추출을 위한  
부가 정보 구성 요소와 형식

Syntax and Semantics of Metadata for Extracting  
Multiple Tiles of Virtual Reality Video



한국정보통신기술협회  
Telecommunications Technology Association

표준초안 검토 위원회 디지털콘텐츠 프로젝트그룹(PG610)

표준안 심의 위원회 소프트웨어/콘텐츠 기술위원회(TC6)

	성명	소속	직위	위원회 및 직위
표준(과제) 제안	류은석	성균관대학교	조교수	PG610 특별위원
	다차원영상기술표준화포럼			
	정종범	성균관대학교	대학원생	-
	이순빈	성균관대학교	대학원생	-
표준 초안 에디터	류은석	성균관대학교	조교수	PG610 특별위원
	김상원	한국전자통신연구원	책임연구원	PG610 간사
	이범렬	한국전자통신연구원	책임연구원	PG610 부의장
	김영민	전자부품연구원	책임연구원	PG610 부의장
	남현우	동덕여자대학교	교수	PG610 의장
사무국 담당	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 약약서 정보는 본 표준의 '부록(지식재산권 약약서 정보)'에 명시하고 있으며, 이후 접수된 지식재산권 약약서는 TTA 웹사이트에서 확인할 수 있습니다. 준용표준인 경우 해당 표준화기구 또는 단체의 웹사이트에서 이를 확인해야 합니다.

본 표준과 관련하여 접수된 약약서 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 : 2020. 12. 10.

# 서 문

## 1 표준의 목적

이 표준의 목적은 몰입형 미디어를 위한 가상현실(virtual reality, VR) 기술을 응용하여 초고화질의 가상현실 영상을 타일 기반 부호화 및 전송 시 다수의 타일을 추출한 후 단일 비트스트림을 생성하는 다중 타일 추출기를 위한 메타데이터를 정의함으로써 사용자 시점 영역에 해당하는 다수의 가상현실 영상 타일 전송 시 연산량 및 처리 시간을 감소하여 가상현실 영상을 위한 저지연 전송 및 렌더링을 제공하는 데 있다.

## 2 주요 내용 요약

이 표준은 사용자의 머리 움직임 추적이 가능한 머리 장착형 영상장치(head-mounted display, HMD)를 통한 타일 기반 가상현실 영상 전송 및 렌더링 시, 타일 개수만큼의 비트스트림을 생성하는 기존 타일 추출기와는 달리 여러 개의 타일을 포함하는 하나의 비트스트림을 생성하는 다중 타일 추출기의 추출 및 메타데이터 기록 기술, 표준 신호 체계 규격(구문과 의미론)을 기술한다. 본 표준은 타일 기반 가상현실 영상 전송 시스템에 적용이 가능하다.

## 3 인용 표준과의 비교

이 표준은 국제 표준단체 MPEG-Immersive 표준 기술 등을 이용하는 시스템을 위한 별도의 독립적인 시그널링 표준으로서, MPEG 비디오 시스템 표준과 직접적인 관련성이 없음.

## Preface

### 1 Purpose

The purpose of this standard is to describe technologies that enable low-latency transmission and rendering for virtual reality (VR) video by extracting multiple tiles and generating a single bitstream, while organizing the required information to the multiple-tile extractor into metadata, thereby reducing the processing time and the computing power requirements.

### 2 Summary

The standard describes the technical and standard specifications (syntax and semantics) about the multiple-tile extractor that extracts multiple tiles and generates a single bitstream in VR video tiled transmission and rendering through head-mounted display (HMD) devices capable of tracking user's movements, while the existing tile extractor extracts the multiple tiles and generates multiple bitstreams. This standard can be applied to VR video tiled streaming system.

### 3 Relationship to Reference Standards

This standard does not directly affect to or influenced by the referenced MPEG-Immersive 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	가상현실 영상 다중 타일 추출을 위한 부가 정보 구성 요소와 형식	2
5.1	다중 타일 추출기 부가 정보 구성 요소와 형식	2
5.2	신호 체계 규격	4
부록 I	필요성	7
부록 II-1	지식재산권 요약서 정보	8
II-2	시험인증 관련 사항	9
II-3	본 표준의 연계(family) 표준	10
II-4	참고 문헌	11
II-5	영문표준 해설서	12
II-6	표준의 이력	13

# 가상현실 영상 다중 타일 추출을 위한 부가 정보 구성 요소와 형식 (Syntax and Semantics of Metadata for Extracting Multiple Tiles of Virtual Reality 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

해당 사항 없음

## 4 약어

DASH	Dynamic Adaptive Streaming over HTTP
EIS	Extraction Information Sets
HEVC	High-Efficiency Video Coding
HM	HEVC Test Model
MCTS	Motion-Constrained Tile Set
MPD	Media Presentation Description
MPEG	Moving Picture Experts Group
NAL	Network Abstraction Layer
PPS	Picture Parameter Set
VPS	Video Parameter Set
SEI	Supplemental Enhancement Information

## SPS Sequence Parameter Set

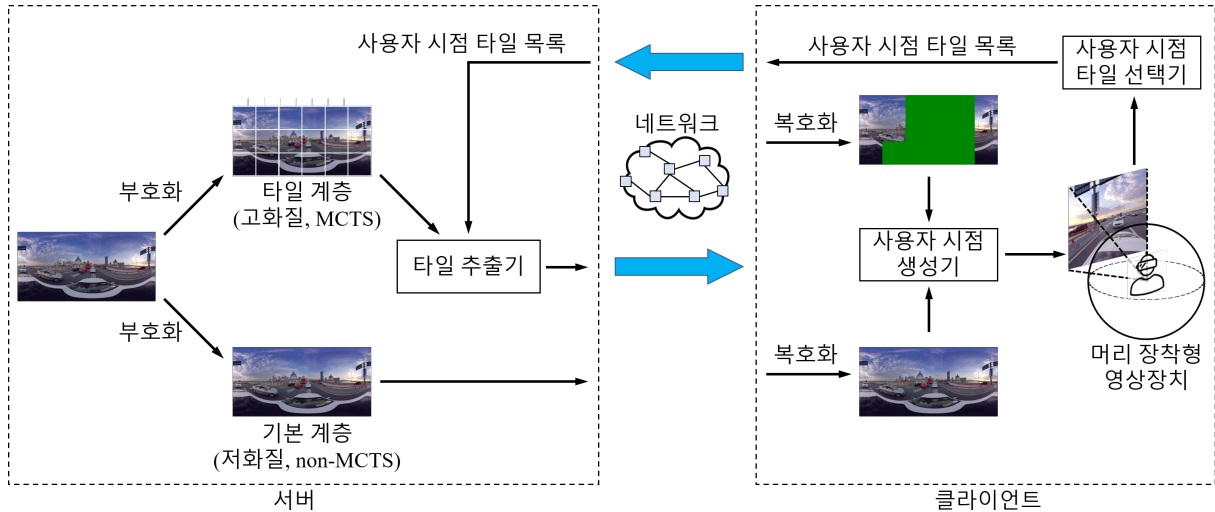
## 5 가상현실 영상 다중 타일 추출을 위한 부가 정보 구성 요소와 형식

## 5.1 다중 타일 추출기 부가 정보 구성 요소와 형식

## 5.1.1 타일 기반 가상현실 영상 스트리밍 시스템

몰입형 가상현실을 위한 360도 영상 스트리밍 시 초고해상도 영상이 필요한 것을 감안하여, 사용자 시점을 판단하여 그에 해당하는 360도 영상의 일부만을 전송하는 기술이 사용될 수 있다. 그러나, 단순히 영상을 픽셀 수준에서 분할하여 부호화 후 비트스트림을 전송하는 것은 서버에 추가적인 연산량과 저장 공간을 요구하므로 바람직하지 않다. 따라서, MPEG에서는 부호화된 비트스트림 수준에서 직사각형 형태의 영역, 즉 타일을 추출할 수 있도록 하는 MCTS 기법과 비트스트림 내에 포함된 타일에 대한 정보를 포함하는 EIS SEI 메시지를 정의하였고, 이는 HEVC에 포함되었다. HEVC 표준을 따르는 참조 소프트웨어인 HM에 포함된 타일 추출기는 타일 한 개당 한 개의 타일 비트스트림을 생성한다. 타일의 개수를 증가시킬 때 타일의 크기가 줄어들어 사용자 시점 영역의 크기가 줄어들어 대역폭이 절감되고, 이 경우 사용자 시점에 해당하는 타일이 여러 개가 됨을 감안했을 때, 클라이언트단에서 복호화해야 할 비트스트림의 개수는 기존 스트리밍 기법 대비 늘어난다. 따라서, 타일 스트리밍 시 단일 타일 추출기보다는 여러 개의 타일을 포함하는 단일 비트스트림을 출력하는 다중 타일 추출기가 더 적은 수의 복호기를 클라이언트에 요구하므로 더 효율적인 스트리밍 환경을 제공할 수 있다.

(그림 5-1)은 사용자 시점에 해당하는 영역을 고화질의 타일로, 전체 영상에 대해서는 저화질의 비트스트림으로 전송하는 타일 기반 360도 영상 스트리밍을 보여준다. 저화질의 비트스트림은 기본 계층으로 나타나고, 고화질의 타일은 타일 계층으로 나타내어진다. 클라이언트는 머리 장착형 영상장치에서 사용자 시점 정보를 사용자 시점 타일 선택기로 전달하고, 이후 사용자 시점 타일 목록을 서버로 전달한다. 이후 서버는 사용자 시점에 해당하는 타일 정보를 다중 타일 추출기로 전달한다. 다중 타일 추출기는 목표 타일들을 추출 후 단일 비트스트림을 생성하여 전송하고, 기본 계층 비트스트림은 이와 상관없이 전송된다. 이후 클라이언트단에서는 타일 계층, 기본 계층에 대응하는 2개의 복호기로 비트스트림들을 복호화하고, 이후 사용자 시점 렌더러로 영상이 전달되면 사용자 시점의 영상이 생성되어 머리 장착형 영상장치에 표현된다. 이렇게 저화질, 고화질 계층을 나누어 전송함으로써 고화질 저지연 몰입형 영상 제공이 가능하다.



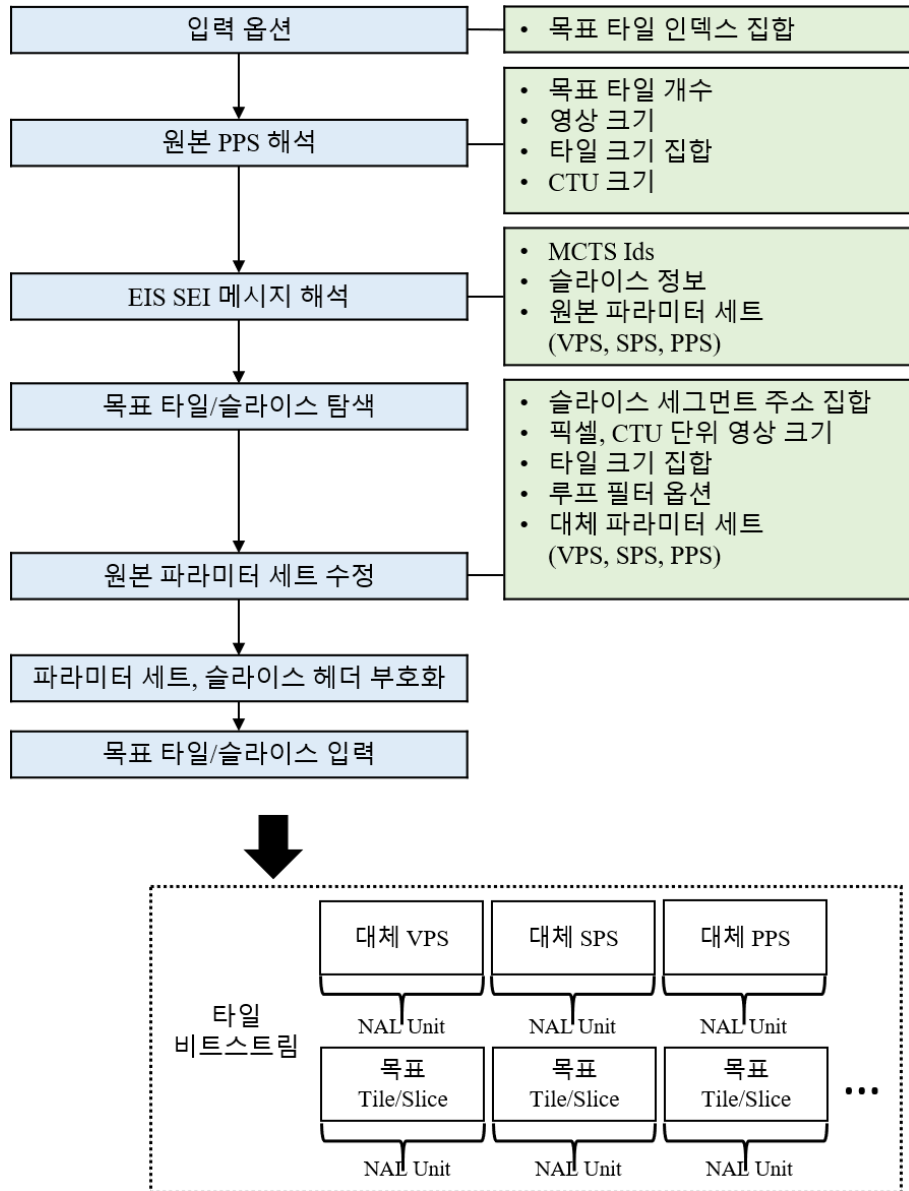
(그림 5-1) 타일 기반 360도 가상현실 영상 스트리밍 구조도

### 5.1.2 다중 타일 추출과 단일 비트스트림 생성

사용자 시점이 변경된 후 해당 시점에 포함되는 타일들이 결정되면, 서버는 전체 영상에 대한 저화질 비트스트림을 전송하면서 동시에 고화질의 사용자 시점 타일들을 추출하여 전송한다. 이 때, 클라이언트는 복호화된 저화질 영상으로 사용자 시점을 렌더링하여 영상을 일시적으로 보여주고 이후 고화질 영상을 머리 장착형 영상장치에 재주사함으로써 사용자에게 몰입감 있는 경험을 제공할 수 있다.

본 표준 기술은 다중 타일 추출을 위해 타일로 분할되어 구성된 비트스트림의 구조 정보를 이용한다. (그림 5-2)는 다중 타일 추출기에서의 여러 개의 타일을 추출하여 단일 비트스트림을 생성하는 과정을 보여준다. 추출해야 할 목표 타일들의 인덱스들이 주어지면, 추출기는 NAL 유닛 단위로 구성된 비트스트림을 분석하여 추출 시 필요한 정보를 저장한다. 비트스트림 복호화 시 필요한 정보들을 가지고 있는 NAL 유닛인 VPS, SPS, PPS 에서 영상 크기, 타일 크기 집합, CTU 크기 등을 가져온다. 이후 EIS SEI 메시지에서 MCTS id, 슬라이스 정보, VPS, SPS, PPS 등을 가져와 저장한 후, 목표 타일들을 포함한 슬라이스들을 탐색 및 추출한다. 추출 완료 후 목표 타일들에 따라 원본 파라미터 세트를 수정하여 슬라이스 세그먼트 주소 집합, 영상 크기, 타일 크기 집합, 루프 필터 옵션 등을 반영한 대체 파라미터 세트를 생성하고, 비트스트림에 삽입하기 위해 슬라이스 헤더와 같이 부호화를 진행한다. 이후, 추출된 슬라이스들을 비트스트림에 입력하면 다중 타일을 포함한 단일 비트스트림이 생성된다. 이렇게 단일 비트스트림 생성 시 여러 개의 비트스트림들을 처리할 때보다 복호화 연산량 및 처리 시간을 감소시킬 수 있다.





(그림 5-2) 다중 타일 추출기의 기능 흐름도

## 5.2 신호 체계 규격

본 표준의 핵심 신호 체계는 비디오 수신 및 렌더링을 담당하는 머리 장착형 영상장치가 360도 가상현실 영상 스트리밍 서버로부터 전달받는 비트스트림에 포함된 타일 크기, 위치 및 개수 정보들을 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 다중 타일 비트스트림 추출 시 SPS에서 수정되어야 할 신호 체계 규격이다. 표에 나온  $u(n)$ 는 통상 프로그래밍 언어에서 부호가 없는(unsigned) 'n' 비트 수를 의미한다.

<표 5-1> 다중 타일 추출 시 SPS에서 변경되는 신호 체계 규격

신호 체계 규격	비트 수
Seq_parameter_set_info {	
...	
pic_width_in_luma_samples	u(16)
pic_height_in_luma_samples	u(16)
...	
}	

다음 <표 5-2>는 다중 타일 비트스트림 추출 시 SPS에서 수정되어야 할 신호 체계에 대한 구문 의미론이다.

<표 5-2> 다중 타일 추출 시 SPS에서 변경되는 정보에 대한 구문 의미론

구문	의미론
pic_width_in_luma_samples	영상의 휘도 너비를 의미함. 부호 없는 16비트의 정수로 표현됨.
pic_height_in_luma_samples	영상의 휘도 높이를 의미함. 부호 없는 16비트의 정수로 표현됨.

다음 <표 5-3>은 다중 타일 비트스트림 추출 시 PPS에서 수정되어야 할 신호 체계 규격이다. 표에 나온 u(n)는 통상 프로그래밍 언어에서 부호가 없는(unsigned) ‘n’ 비트 수를 의미한다.

<표 5-3> 다중 타일 추출 시 PPS에서 변경되는 신호 체계 규격

신호 체계 규격	비트 수
Pic_parameter_set_info {	
...	
tiles_enabled	u(1)
if (tiles_enabled) {	
num_tile_in_columns	u(8)
num_tile_in_rows	u(8)
uniform_spacing	u(1)

신호 체계 규격	비트 수
loop_filter_across_tiles_enabled	u(1)
}	
loop_filter_across_slices_enabled	u(1)
...	
}	

다음 <표 5-4>는 다중 타일 비트스트림 추출 시 PPS에서 수정되어야 할 신호 체계에 대한 구문 의미론이다.

<표 5-4> 다중 타일 추출 시 PPS에서 변경되는 정보에 대한 구문 의미론

구문	의미론
tiles_enabled	영상이 타일로 분할되었는지를 의미함. 부호 없는 1비트의 정수로 표현됨.
num_tile_in_columns	영상 내 타일들의 열 개수를 의미. 부호 없는 8비트의 정수로 표현됨.
num_tile_in_rows	영상 내 타일들의 행 개수를 의미. 부호 없는 8비트의 정수로 표현됨.
uniform_spacing	영상 내 타일들이 균등한 크기로 분할되었는지 의미. 부호 없는 1비트의 정수로 표현됨.
loop_filter_across_tiles_enabled	영상 내 타일 간 루프 필터가 활성화되었는지 의미. 부호 없는 1비트의 정수로 표현됨.
loop_filter_across_slices_enabled	영상 내 슬라이스 간 루프 필터가 활성화되었는지 의미. 부호 없는 1비트의 정수로 표현됨.

## 부 록 I

(본 부록은 표준을 보충하기 위한 내용으로 표준의 일부는 아님)

### 필요성

#### 1.1 본 표준의 필요성

최근 가상현실에 대한 관심이 높아지면서, 가상현실 장비 및 기술이 활발히 개발되고 있다. 머리 장착형 영상장치를 통해 360도 영상을 렌더링 및 표현하는 기술, 360도 영상을 부호화 및 전송하는 등의 기술들이 발전하고 있다. 사용자의 quality of experience(QoE)를 높이고 불편감을 경감시키기 위해 ultra high-definition(UHD)급 이상의 초고화질 360도 영상이 요구되는데, 이는 전송 시 높은 대역폭을 요구한다. 한편, 기존의 2D perspective 영상 렌더링과는 달리, 머리 장착형 영상장치를 통한 360도 가상현실 영상 렌더링 시 사용자의 field of view(FoV)만큼, 즉 영상의 일부만이 화면에 출력된다. 몰입형 가상현실을 위한 360도 영상 스트리밍 시 초고해상도 영상이 필요한데 머리 장착형 영상장치에 표시되는 영상은 전체 영상의 일부인 것을 감안하여, 사용자 시점을 판단하여 그에 해당하는 360도 영상의 일부만을 전송하는 기술이 사용될 수 있다. 따라서, 360도 영상에 대한 효율적인 부호화 및 전송 기법들이 요구되었고, 그 예로 영상을 분할하여 사용자 관심 영역만 전송하는 기법이 연구되고 있다.

ISO/IEC 산하 국제 표준화 단체인 moving picture experts group(MPEG)에서 논의 중인 omnidirectional media format(OMAF)는 high-efficiency video coding(HEVC)를 이용하여 사용자 시점 기반 전방위 영상 부호화 및 전송 기법의 하나로 motion-constrained tile set(MCTS) 부호화 및 타일 추출, 전송 기법을 정의한다. MCTS를 통해 비트스트림 단계에서 하나 이상의 직사각형 형태의 타일로 분할된 비트스트림에서 독립적으로 복호화가 가능한 타일을 추출 가능하고, HEVC의 참조 소프트웨어인 HEVC test model(HM)에 타일 추출기의 구현물이 포함되어 있다. 하지만, HM에서 구현된 타일 추출기는 각각의 타일에 대응하는 여러 개의 비트스트림들을 생성한다. 사용자 시점 타일의 개수가 늘어나면 클라이언트에 요구되는 복호기의 수가 증가하여 자원 소모 및 처리 시간이 늘어나는 단점이 있다. 따라서, 본 표준은 여러 개의 타일을 포함하는 단일 비트스트림을 생성하기 위한 방법 및 신호 체계 규격을 정의 및 기술한다. 본 표준을 통해 대역폭을 절약하기 위해 타일의 크기를 작게 하여 많은 양의 타일을 전송할 때에도 클라이언트에 단일 복호기만을 요구하여 효율적인 스트리밍을 구현할 수 있다.

## 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.

#### II-1.1 지식재산권 확약서(1)

- 발명의 명칭: 타일 기반 360도 영상 전송 시스템 및 방법
- 권리자의 성명: 성균관대학교 산학협력단
- 출원 번호: 10-2020-0086861
- 출원 연월일: 2020년 7월 14일
- 실시조건:
- 확약서 접수일: 2020년 7월 14일

## 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

해당 사항 없음.

## 부 록 II-4

(본 부록은 표준을 보충하기 위한 내용으로 표준의 일부는 아님)

### 참고 문헌

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	2020.12.10	제정 TTAK.KO-KO.1258	-	디지털콘텐츠 프로젝트그룹 (PG610)