

TTA Standard

정보통신단체표준(국문표준)

TTAK.KO-10.1393

제정일: 2022년 12월 07일

사용자 시점 기반 가상현실 영상 타일
병합을 위한 부가정보 구성 요소 및
형식

Syntax and Semantics of Metadata for Tile Merging 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	부교수	PG610 특별위원
	정종범	성균관대학교	대학원생	-
	이순빈	성균관대학교	대학원생	-
다차원영상기술표준화포럼				
표준 초안 에디터	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
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 약삭서 정보는 본 표준의 '부록(지식재산권 약삭서 정보)'에 명시하고 있으며, 이후 접수된 지식재산권 약삭서는 TTA 웹사이트에서 확인할 수 있습니다. 준용표준인 경우 해당 표준화기구 또는 단체의 웹사이트에서 이를 확인해야 합니다.

본 표준과 관련하여 접수된 약삭서 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 가상 현실 영상을 타일 기반 부호화 및 전송 시 타일 고속 추출 및 병합을 통한 단일 비트스트림 생성 및 사용자 시점 영역 기반 품질 할당에 필요한 메타데이터를 정의함으로써 가상현실 영상 전송 시 연산량, 처리 시간을 감소시키고 대역폭 대비 품질을 향상시켜 가상현실 영상을 위한 저지연 전송 및 고품질 렌더링을 제공하는 데 있다.

2 주요 내용 요약

이 표준은 사용자의 머리 움직임 추적이 가능한 머리 장착형 영상장치(head-mounted display, HMD) 및 분산 처리가 가능한 서버를 통한 타일 기반 가상현실 영상 전송 및 렌더링 시, 압축된 영역에서 여러 개의 타일들의 타일/슬라이스 수준 단일 비트스트림으로의 병합 및 선택적 스트리밍을 위한 표준 신호 체계 규격(구문과 의미론)을 기술한다. 본 표준은 타일 기반 가상현실 영상 전송 시스템에 적용이 가능하다. 이 표준은 국제 표준 단체 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 low-latency transmission and rendering for virtual reality (VR) video by extracting and merging tiles with low-latency to generate a single bitstream, while organizing the required information into metadata considering the viewport-adaptive quality allocation, thereby reducing the processing time and the computing power requirements and increasing the quality considering the bandwidth.

2 Summary

This standard describes the technical and standard specifications (syntax and semantics) about the compressed domain merging of multiple tiles/slices into a single bitstream and selective streaming in VR video tile-based transmission and rendering through a head-mounted display (HMD) devices capable of tracking user's movements and a server which supports distributed processing. This standard can be applied to the tile-based 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 사용자 시점 기반 가상현실 영상 타일 병합을 위한 부가정보 구성 요소 및 형식	2
5.1 타일 병합 부가 정보 구성 요소 및 형식	2
5.2 신호 체계 규격	5
부록 I 필요성	8
부록 II-1 지식재산권 확약서 정보	9
II-2 시험인증 관련 사항	10
II-3 본 표준의 연계(family) 표준	11
II-4 참고 문헌	12
II-5 영문표준 해설서	13
II-6 표준의 이력	14

사용자 시점 기반 가상현실 영상 타일 병합을 위한 부가정보 구성 요소 및 형식 (Syntax and Semantics of Metadata for Tile Merging 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

3.1 Exponential Golomb Code

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

4 약어

AVC	Advanced Video Coding
CTU	Coding Tree Unit
DASH	Dynamic Adaptive Streaming over HTTP
EIS	Extraction Information Sets
HEVC	High-Efficiency Video Coding
MCTS	Motion-Constrained Tile Set
MPD	Media Presentation Description
MPEG	Moving Picture Experts Group

NAL	Network Abstraction Layer
PPS	Picture Parameter Set
QP	Quantization Parameter
SEI	Supplemental Enhancement Information
SPS	Sequence Parameter Set
VPS	Video Parameter Set
VVC	Versatile Video Coding

5 사용자 시점 기반 가상현실 영상 타일 병합을 위한 부가정보 구성 요소 및 형식

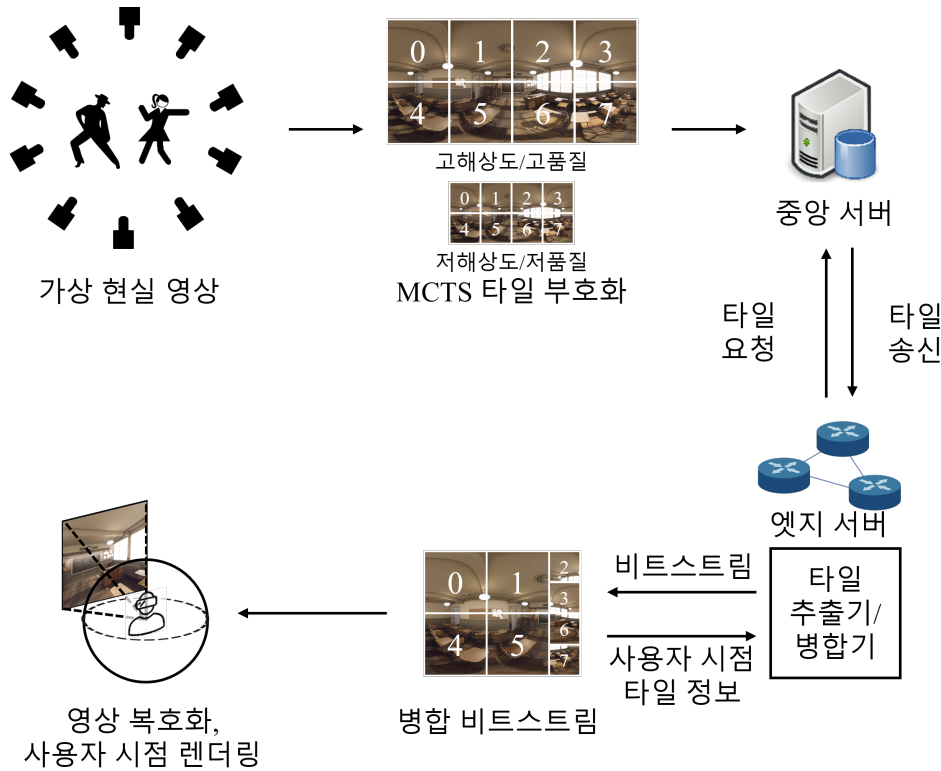
5.1 타일 병합 부가 정보 구성 요소 및 형식

5.1.1 가상현실 영상 타일 스트리밍 시스템

머리 장착형 영상장치를 통해 몰입형 가상현실을 위한 360도 영상 스트리밍 시 영상의 일부만이 렌더링되는 것에 착안하여, 영상에서 사용자 시점에 해당하는 영역만을 추출하여 전송하는 기술이 사용될 수 있다. 단순히 영상을 픽셀 수준에서 분할하여 부호화 후 비트스트림을 전송하는 것은 종래의 영상 압축 도구인 AVC 등을 통해 가능하나, 서버에 추가적인 연산량과 저장 공간을 요구하는 단점이 있다. 최근 널리 사용되는 영상 압축 도구인 HEVC에서는 부호화된 비트스트림 수준에서 직사각형 형태의 영역인 타일 (tile) 을 분할 할 수 있도록 하는 MCTS 기법과 비트스트림 내에 포함된 타일에 대한 정보를 포함하는 EIS SEI 메시지를 정의하여, 압축된 영역에서 각 타일의 독립적인 추출 및 복호화가 가능하다. 추출된 타일 비트스트림들을 HEVC 및 차세대 영상 압축 도구인 VVC 표준에 맞게 압축된 영역에서 병합 후 단일 비트스트림을 생성하는 타일 병합기를 통해 클라이언트에 요구되는 복호기의 수를 줄일 수 있고 지연 시간을 절감할 수 있다. 부호화 및 타일 추출은 중앙 서버에서, 타일 병합은 엣지 (edge) 서버에서 분산 처리하여 적응적인 스트리밍 환경을 제공할 수 있다.

(그림 5-1)은 사용자 시점 기반 가상현실 영상 타일 스트리밍 구조도를 나타낸다. 고화질의 사용자 시점 영상 전송을 위해 중앙 서버는 고해상도/고품질 MCTS 비트스트림을 MCTS 타일 부호화를 통해 생성한다. 사용자 시점 변경 시 사용자 시점 고해상도/고품질 영상의 부재로 인한 사용자 만족도 하락을 방지하기 위해 중앙 서버는 추가적으로 저해상도/저품질 MCTS 비트스트림을 생성한다. 중앙 서버는 엣지 서버에서 요청한 고해상도/고품질 MCTS 비트스트림과 저해상도/저품질 MCTS 비트스트림을 엣지 서버로 전송한다. 엣지 서버에 위치한 타일 추출기/병합기는 MCTS 비트스트림들에서 각 타일들을 추출하고, 클라이언트에서 전송받은 사용자 시점 타일 인덱스에 해당하는 고해상도/고품질 타일 비트스트림들과 그 외 영역을 담당하는 저해상도/저품질 타일 비트스트림들을

HEVC 및 VVC의 타일 규칙에 맞게 병합한다. 병합 과정은 병합된 비트스트림의 픽처 (picture) 해상도를 최소화하도록 타일들을 배열하며 진행된다. 이후 클라이언트는 엣지 서버로부터 전송받은 병합 비트스트림을 복호화하고, 사용자 시점을 생성하여 머리 장착형 영상장치에 표현한다.



(그림 5-1) 360도 가상현실 영상 사용자 시점 기반 타일 스트리밍 구조도

5.1.2 타일/슬라이스 수준 비트스트림 병합

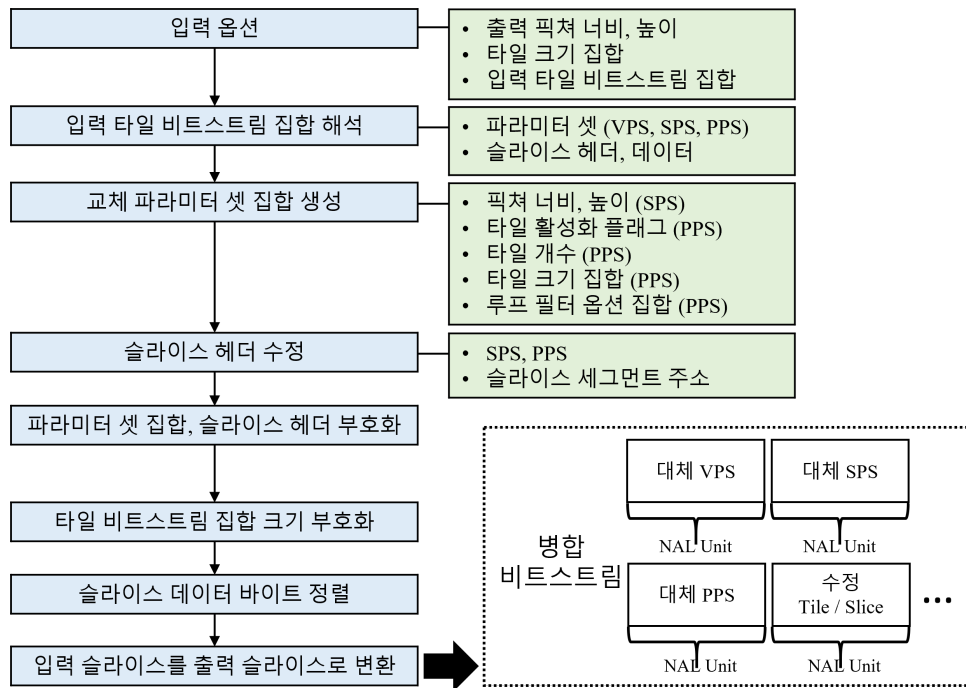
사용자 시점 변경 후 해당 시점에 포함되는 타일들이 계산되면, 엣지 서버는 사용자 시점 고해상도/고품질 타일 비트스트림들과 그 외의 영역을 담당하는 저해상도/저품질 타일 비트스트림들을 병합하여 클라이언트로 전송한다. HEVC 및 VVC의 타일 규칙에 따르면, 한 타일 내에 여러 슬라이스 (slice) 들이 수직 방향으로 배열될 수 있다. 상기 규칙을 따르면서 다양한 해상도를 가진 타일 비트스트림들을 병합하여 출력 비트스트림의 해상도를 감소시키기 위해, 엣지 서버에 위치한 타일 병합기는 고해상도 타일 기준으로 타일의 경계를 정의하고 해당 타일의 크기에 맞게 여러 개의 저해상도 타일들을 슬라이스 단위로 수직 방향으로 배열한다. 다양한 품질을 가진 타일 비트스트림들을 병합하기 위해서 엣지 서버에 위치한 타일 병합기는 첫 타일을 기준으로 품질을 정의하는 QP 값을 파라미터 셋에 저장하고, 첫 타일을 제외한 타일들에 대해서는 저장된 QP 값을 기준으로 차분값만을 저장한다. (그림 5-2)는 타일/슬라이스 수준 비트스트림 병합 예시를 나타낸다. 저해상도/저품질 MCTS 비트스트림은 고해상도/고품질 MCTS 비트스트림의 절반에

해당하는 너비와 높이를 가진다. 사용자 시점 타일 인덱스가 0, 1, 4, 5일 때, 고해상도/고품질 타일 0, 1, 4, 5를 기준으로 출력 비트스트림의 타일 행과 열의 크기가 결정된다. 이후 저해상도/저품질 타일 2, 3은 고해상도/고품질 타일 1의 좌측에, 저해상도/저품질 타일 6, 7은 고해상도/고품질 타일 5의 좌측에 위치하도록 슬라이스 단위로 삽입된다. 이후 복호기를 통해 병합 비트스트림을 복호화하면 (그림 5-2)와 도시된 것과 같이 타일/슬라이스 경계에서 문제 없이 복호화됨을 확인할 수 있다.



(그림 5-2) 타일/슬라이스 수준 비트스트림 병합 예시

본 표준 기술은 타일 병합을 위해 타일로 분할되어 구성된 비트스트림의 구조 정보를 이용한다. (그림 5-3)은 타일 병합기의 기능 흐름을 보여준다. 입력 옵션으로 출력 픽처의 너비, 높이, 타일 크기 집합, 입력 타일 비트스트림들의 이름이 주어진다. 병합기는 타일 비트스트림들을 NAL 유닛 단위로 해석하여 VPS, SPS, PPS 등의 파라미터 셋을 확보하고, 각 타일들을 포함하는 슬라이스 NAL 유닛은 헤더만을 해석하여 저장하고 실질적으로 부호화된 타일 정보들을 포함하는 데이터는 해석하지 않아 처리 시간을 절약한다. 병합 비트스트림에 맞는 정보들을 교체 파라미터 셋에 반영하기 위해 픽처 너비, 높이를 SPS에, 타일 활성화 플래그, 타일 개수, 타일 크기 집합, 루프 필터 (loop filter) 옵션 집합을 PPS에 저장한다. 이후 슬라이스 헤더에서는 교체 파라미터 셋에 해당하는 SPS, PPS 및 해당 슬라이스가 포함하는 타일에 대한 병합 비트스트림에서의 슬라이스 세그먼트 주소를 반영한다. 파라미터 셋 집합, 슬라이스 헤더를 부호화한 후 타일 비트스트림들의 크기들을 각각 부호화하고, 슬라이스 데이터를 바이트 단위로 정렬한다. 이후, 입력 슬라이스들을 출력 슬라이스로 변환 후 병합 비트스트림을 출력한다.



(그림 5-3) 타일 병합기 기능 흐름도

5.2 신호 체계 규격

본 표준의 핵심 신호 체계는 타일 병합기가 360도 가상현실 영상 스트리밍 중앙 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 타일 병합 시 PPS에서 수정되어야 할 신호 체계 규격이다. 표에 나온 $se(v)$ 는 통상 프로그래밍 언어에서 부호가 있는(signed) exponential golomb code로 부호화된 정수를 의미한다.

<표 5-1> 타일 병합 시 PPS에서 변경되는 신호 체계 규격

신호 체계 규격	비트 수
Seq_parameter_set_info {	
...	
init_qp_minus26	$se(v)$
...	
}	

다음 <표 5-2>는 타일 병합 시 PPS에서 수정되어야 할 신호 체계에 대한 구문 의미론이다.

<표 5-2> 타일 병합 시 PPS에서 변경되는 정보에 대한 구문 의미론

구문	의미론
init_qp_minus26	영상의 QP에서 26을 뺀 값을 의미함. 부호 있는 exponential golomb code로 부호화된 정수로 표현됨.

다음 <표 5-3>은 타일 병합 시 슬라이스 세그먼트 (segment) 헤더에서 수정되어야 할 신호 체계 규격이다. 표에 나온 u(n)은 통상 프로그래밍 언어에서 부호가 없는 (unsigned) 'n' 비트 수를 의미한다.

<표 5-3> 타일 병합 시 슬라이스 세그먼트 헤더에서 변경되는 신호 체계 규격

신호 체계 규격	비트 수
Slice_segment_header_set_info {	
...	
first_slice_segment_in_pic	u(1)
...	
if (!is_first_slice_segment_in_pic) {	
slice_segment_address	u(v)
}	
...	
slice_qp_delta	se(v)
...	
}	

다음 <표 5-4>는 타일 병합 시 슬라이스 세그먼트 헤더에서 수정되어야 할 신호 체계에 대한 구문 의미론이다.

<표 5-4> 타일 병합 시 슬라이스 세그먼트 헤더에서 변경되는 정보에 대한 구문 의미론

구문	의미론
first_slice_segment_in_pic	슬라이스가 픽처 내에서 첫 번째 슬라이스인지를 의미함. 부호 없는 1비트의 정수로 표현됨.

구문	의미론
slice_segment_address	픽처 내 슬라이스의 CTU 단위 시작 주소를 의미함. 부호 없는 정수로 표현됨.
slice_qp_delta	슬라이스의 QP 차분값을 의미함. 부호 있는 exponential golomb code로 부호화된 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. 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 비트스트림을 중앙 서버에서 생성하고, 엣지 서버에서 타일 추출 및 타일들을 병합하여 단일 비트스트림을 생성하는 타일 병합기를 설치하면 분산 처리를 통해 지연 시간을 절약가능하다. 따라서, 본 표준은 타일들을 병합하여 단일 비트스트림을 생성하기 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-10.1258, 가상현실 영상 다중 타일 추출을 위한 부가 정보 구성 요소와 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 11-5

(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-10.1393	-	디지털콘텐츠 프로젝트그룹 (PG610)