

# MPEG Media Transport 환경 내 UWV 미디어 공간정보 시그널링 방안

김준식, 강전호, 이의상, 김규현  
경희대학교

junsik@khu.ac.kr, gaonam@khu.ac.kr, euisang91@khu.ac.kr, kyuheonkim@khu.ac.kr

## Signaling Method for Spatial Information of UWV media under MPEG Media Transport Environment

Junsik Kim, Jeonho Kang, Euisang Lee, Kyuheon Kim  
Kyunghee University

### 요 약

UWV(Ultra Wide View) 미디어 전송 시 서버의 부하를 줄이고 단말기의 자원 소모를 줄이기 위해서는 사용자의 관심 영역만을 제공해야 한다. 즉, 사용자가 원하는 범위의 영상들을 선택할 수 있도록 영상들의 공간적 위치정보를 제공하는 기술이 필요하다. 본 논문에서는 MMT(MPEG Media Transport)의 CI(Composition Information) 문서를 이용하여 UWV 를 구성하는 미디어들의 공간적 위치정보를 전송하고, 사용자에게 관심 영역의 미디어들을 제공하는 기법을 제안한다. 이를 위해 서버가 사용자에게 미디어들의 화면 구성에 대한 공간적 정보와 재생에 필요한 시간적 정보를 MMT CI 문서에서 제공한다. 그리고 사용자의 관심 영역을 서버로 전송하기 위한 Signaling Message 를 정의한다. 서버는 본 논문의 기법을 통해 MMT 를 그대로 사용하면서 사용자에게 관심 영역의 미디어들을 제공할 수 있다.

### 1. 서론

최근 영상 기술은 빠르게 발전하고 있다. 이는 다양한 미디어들의 생산을 촉진하여 파노라마 영상, 360 Video, VR 등 다양한 종류의 미디어의 생산과 연결되었다. 특히 최근의 영상기술들은 넓은 시야각과 선명한 화질을 통해 실제와 같은 현실감을 제공하는 UWV(Ultra Wide View)의 경향성을 띄고 있다.[1] 본 논문에서는 상기한 경향성을 가진 미디어들을 UWV 미디어로 총칭한다. UWV 미디어들은 사용자에게 현장감을 주며, 넓은 시야각을 통해 다양한 정보를 제공한다는 장점을 갖고 있다. 하지만 현재 UWV 미디어의 상용화에는 여러 문제점이 존재한다. 먼저 UWV 미디어 생산에는 다양한 시점의 미디어를 Stitching 하는 작업이 필요하다. 하지만 Stitching 작업은 상당한 계산 작업이 요구되어, 실시간 전송 시 서버의 부하가 급증하는 결과를 가져온다. 그리고 사용자는 비 관심 영역의 미디어를 전송 받을 때에도 큰 자원을 소모해야 한다는 단점도 존재한다. 따라서 UWV 를 구성하는 여러 미디어 중 사용자의 관심 영역에 존재하는 미디어들만을 선별적으로 전송하는 시스템이 필요하다. 본 논문에서는 전송 시 UWV 미디어를 구성하는 미디어의 공간적 위치 정보를 전송하고, 사용자는 signaling message 를 통해 관심 영역을 서버로 전달하는 기법을 제안하여 상기 문제를 해결하고자 한다. 미디어들의 공간적 위치 정보를 Spatial Adjacency Matrix 로 명명하고[2], 인접 미디어간 행렬의 형태로 전송한다. 전송 시에는 MMT(MPEG Media Transport)의 CI(Composition Information)를 사용하여, Spatial Adjacency Matrix 를 MMT CI 에 담아 전송하는 기법을 제안한다.

제안 기술에 관한 본 논문의 구성은 다음과 같다. 2 장에서는 기반 기술인 MMT CI 에 대해 설명하고, 3

장에서는 Spatial Adjacency Matrix 를 MMT CI 에 담은 기법과 사용자의 관심 영역을 서버에 전달하는 Signaling Message 에 대해 정의한다. 논문 제안 사항들을 4 장에서 실험 및 결과를 통해 밝히며, 마지막으로 5 장에서 본 논문에 대한 결론을 맺는다.

### 2. MMT CI(Composition Information)

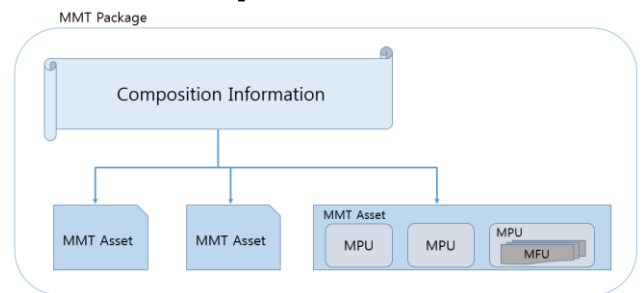


그림 1. MMT Package 내에서 MMT CI 의 역할

MMT 는 IP 기반의 미디어 전송 서비스를 제공하는 MPEG 의 방송 시스템 표준이다. MMT 는 Package 라는 논리적 콘텐츠 집합으로 구성되어 있고, 각각의 Package 는 여러 개의 Assets 으로 구성되어 있다. Asset 은 전송 및 재생과 저장의 최소단위인 MPU(Media Processing Unit)들로 이루어진 동일한 미디어 특성이 부여되는 가장 큰 단위를 말한다. 이 MPU 들은 frame 단위의 MFU(Media Fragment Unit)들로 구성되어 있다. 하나의 MMT Package 는 <그림 1>에서 보이는 바와 같이 CI 문서를 통해 Asset 간의 관계와 MPU 의 시간 및 공간적 관계 정보를 제공해야 한다.[3]

MMT 의 Composition Layer 는 CI 문서를 통해 MMT

단일 Package 내 Assets 또는 MPU 간의 공간 관계에 대한 정보와 시간 정보를 제공하는 계층으로, Composition Layer 는 마크업 언어 중 하나인 XML 과 HTML5 기술을 이용한다.

HTML5(Hyper Text Markup Language)는 웹문서 작성을 목표로 개발된 마크업 언어로, 웹페이지 상 다양한 소스의 공간구성이 가능하다는 장점이 있다. Composition Layer 에서는 HTML5 를 이용해 미디어 요소간의 초기 공간적 관계 정보를 제공한다.

XML(Extensible Markup Language)은 W3C 에서 개발된 다목적 마크업 언어이다. 사람과 기계가 모두 읽을 수 있는 형식을 규정하고 있는데, 이는 HTML5 의 복잡성에서 오는 한계를 극복하고 시스템간의 데이터 전송을 쉽게 하는데 목적이 있다. HTML5 문서는 미디어 요소간의 초기 공간 정보만을 제공하기 때문에, Composition Layer 에서는 XML 을 이용해 단일 Package 내의 Asset 또는 MPU 간의 시간적 관계에 대한 정보를 제공한다.

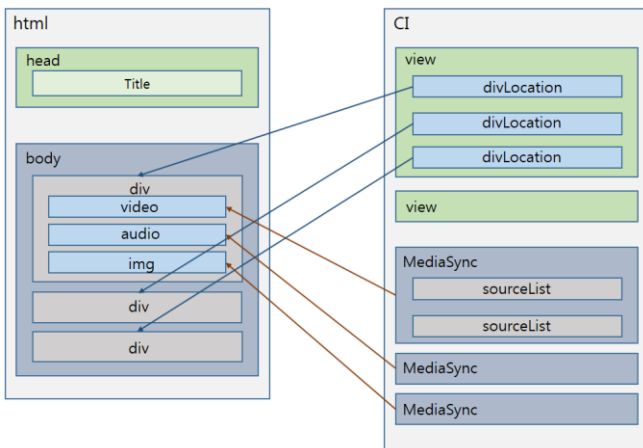


그림 2. CI 문서 구조와 이와 관계된 HTML5 문서

<그림 2>의 좌측은 HTML5 문서, 우측은 XML 로 작성된 CI 문서를 나타내고 있다. 좌측 HTML5 문서는 <head>와 <body>로 구성되어 있다. <head>는 제공하는 Asset 의 전체 특징을 정의하며, <body>는 <div>를 통해 분할된 각각의 화면 정보를 제공한다. 이 <div>는 <video>, <audio>, <image> 등의 미디어 요소를 갖고, 각각의 미디어 요소들은 초기 화면 구성 정보를 <Attribute>에 갖는다. 각각의 멀티 소스들은 XML 문서 내의 <MediaSync>와 연결되며, <MediaSync> 내의 시간 정보를 통해 미디어 요소들의 시작 및 종료 시간 정보를 확인할 수 있다. 분할된 화면들의 공간적 변화는 CI 문서 내의 <view>를 통해 제공된다. <view>는 <div>와 연결되는 <divLocation>들을 통해 공간 변화 정보와 시작 및 종료 시간 정보를 함께 제공하고 있다.[4]

2 장에서는 화면의 초기 공간 정보와 미디어 요소들의 시간 정보가 기술 되는 MMT CI 에 대해 살펴보았다. 3 장에서는 상기한 MMT CI 를 기반으로 Spatial Adjacency Matrix 를 전송하는 방안에 대해 살펴보도록 한다.

### 3. 제안 기술 소개

#### 가. MMT CI 기반 Spatial Adjacency Matrix 정보 기술 방안

본 논문에서는 HTML5 와 CI 문서의 공간적, 시간적

정보를 활용하여, MMT CI 에 UWV 미디어를 구성하는 미디어들의 Spatial Adjacency Matrix 를 기술하고 각 미디어들의 재생 시간을 기술하는 개념을 도입한다. 상기 정보들을 통해 단말기는 UWV 미디어를 이루는 각 미디어들의 공간적 구성 형태를 알 수 있다. 또한 단말기는 원하는 범위의 관심 영역 영상만을 받아 볼 수 있다.

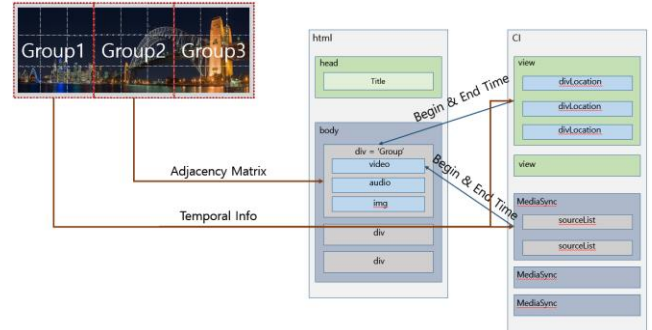


그림 3. 'Group' 생성 예시 및 재생 시간 정보 기술 방법

먼저 UWV 미디어의 영역을 <그림 3>에서 보이는 바와 같이 나눈다. <그림 4>의 경우 총 3 개의 영역으로 나뉘어 있다. 하나의 영역은 MMT 의 Asset 으로 전송되며, div 단위에서는 'Group' 으로 명명한다. 하나의 'Group' 에는 UWV 미디어에서 동일한 범위를 가지는 미디어들을 포함시킨다. <그림 4>의 경우 하나의 'Group' 에 9 개의 미디어가 포함되어 있음을 알 수 있다. 이후 사용자에게 제공할 영역에 맞추어 <view>를 생성한다. 이 <view>들은 사용자가 원하는 정보를 제공하기 위한 모든 경우의 수를 포함하여 제공된다. 즉, 단말기는 설정된 <view>를 기반으로 관심 영역을 선택한다.

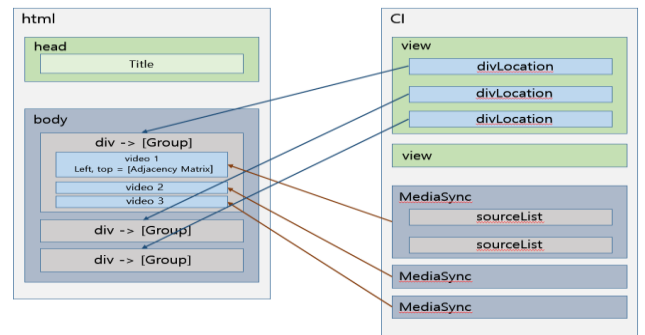


그림 4. Spatial Adjacency Matrix 기술 방법

HTML5 로 작성된 CI 문서에서는 <body> 내의 <div>가 하나의 'Group' 을 갖고 <div>내의 <id>가 'Group' 의 이름을 나타낸다. <div>는 Group 을 구성하는 미디어 요소인 <video>, <audio>, <image> 정보를 포함하며, 이들은 <그림 4>에 보여지듯이 Attribute 인 <left>와 <top>을 통해 각각 자신의 Spatial Adjacency Matrix 정보를 제공한다. HTML5 문서를 통해 사용자는 전체 UWV 를 구성하는 'Group' 정보와 Group 을 구성하는 미디어들의 Spatial Adjacency Matrix 정보를 획득할 수 있다.

각 <div>는 CI 문서의 여러 <view> 와 연결된다. 각 <view>의 <divLocation>들은 필요한 <div>와 연결되어 <view>의 시작 및 종료 시간과 위치정보를 함께 제공한다. 그리고 <MediaSync>를 통해 각각의 미디어 요소의 시작 및 종료 시간에 대한 정보를 얻을 수 있다.

### 나. ROI(Region Of Interest) Message

상기한 기법을 통해 사용자는 UWV 미디어의 구성 요소들을 파악할 수 있으며, Spatial Adjacency Matrix 를 이용해 미디어 요소들의 공간적 정보를 얻을 수 있다. 이제 사용자는 CI 를 통해 얻는 정보로 관심 영역을 서버에게 요구할 수 있다. 본 논문에서는 이러한 작업을 위해 Signaling Message 를 정의한다.

Syntax	Value	No. of bits	Mnemonic
ROI_message () { message_id version length extension { view_id_length if(view_id_length > 0){ for(i=0;i<N1;i++){ view_id [i] } } else{ number_of_media for(i=0;i<N2;i++){ media_id_length for(j=0;j<N3;j++){ media_id [j] } } } } }	N1	16	uimsbf
		32	uimsbf
		32	char
	N2	16	uimsbf
	N3	16	uimsbf
		32	char

그림 5. Syntax of ROI Message

<그림 5>는 ROI message 로 명명한 signaling message 의 Syntax 이다. 먼저 사용자는 ROI message 의 <view\_id>를 이용해 서버가 MMT CI 를 통해 제공한 <view> 중 관심 영역의 <view>를 요구할 수 있다. 또한 사용자의 관심 영역이 제공된 <view> 내에 존재하지 않을 경우 <number\_of\_media>와 <media\_id>를 이용하여 서버에게 관심 영역의 미디어들을 요구할 수 있다. <media\_id>는 CI 문서의 <MediaSync> 내 <refId>값을 그대로 사용한다.

#### Semantics

- message\_id - indicates the identifier of the ROI messages.
- version - indicates the version of the ROI messages. The newer version obsoletes the information in any older version.
- length - a 32-bit field for conveying the length of the ROI message in bytes, counting from the beginning of the next field to the last byte of the ROI message. The value '0' is not valid for this field
- view\_id\_length - indicates the view\_id length in chars. The negative value is not valid for this field
- view\_id - indicates the view\_id as defined in MMT CI document.
- number of media - indicates the number of media. The value '0' is not valid for this field
- media\_id\_length - indicates the media\_id length in chars. The negative value is not valid for this field
- media\_id - indicates the media\_id as equal to refid in MediaSync of MMT CI document.

그림 6. Semantics of ROI Message

Syntax 의 변수들은 <그림 6>의 Semantics 를 통해 구체적으로 정의된다. <view\_id\_length>값이 '0'이 아니라면 <view>를 요구하는 ROI message 임을 알 수 있으며, <view\_id\_length>값이 '0'이라면 <number\_of\_media> 값이 존재하고, 이 값은 '0'을 가질 수 없다. 따라서 <view> 혹은 미디어 중 한가지를 반드시 요구하는 ROI message 의 특징을 확인할 수 있다.

3 장에서 제안한 기술을 통해 사용자는 서버가 전송한 Spatial Adjacency Matrix 에서 UWV 미디어의 구성 정보를 얻게 되고, Spatial Adjacency Matrix 를 기반으로 자신의 관심영역을 서버에게 요구할 수 있게 된다. 4 장에서는 3 장에서

제안한 기법을 직접 구현하고 실험하여 검증할 것이다.

## 4. 구현 및 실험 결과

본 논문에서 제안한 기술은 서버가 MMT CI 를 기반으로 사용자에게 Spatial Adjacency Matrix 를 전송하여 UWV 미디어들이 어떻게 구성되어 있는지를 밝히며, 사용자는 ROI Message 를 통해 서버에게 자신이 원하는 관심 영역 정보를 전달하는 방안이다. 검증은 총 7 개의 미디어를 사용한 UWV 미디어를 통해 이루어진다. 서버는 미디어들을 2 개의 'Group'으로 나누어 총 3 가지의 <view>를 구성하고, 사용자는 ROI message 를 통해 관심 영역을 요구하는 시나리오로 제안 기술을 검증한다.

검증 시나리오를 위해 먼저 MMT CI 의 HTML5 문서와 CI 문서를 구현하였다.

```
<!DOCTYPE html>
<html xmlns="http://www.w3.org/1999/xhtml">
  <head>
    <title>Sample</title>
  </head>
  <body>
    <div id="Group1" style="position:absolute; width:1487px; height:855px">
      
      
      
    </div>
    <div id="Group2" style="position:absolute; width:1913px; height:719px">
      
      
      
      
    </div>
  </body>
</html>
```

그림 7. HTML5 문서를 이용한 제안 기술 구현

<그림 7>은 시나리오의 HTML5 문서이다. 서버가 제공하는 2 개의 'Group'을 각각 'Group1'과 'Group2'로 명명 하였다. 'Group1'은 3 개의 미디어로 구성되고, 'Group2'는 4 개의 미디어로 구성되는 것을 확인할 수 있다. 또한 이들은 Attribute 인 <left>와 <top>에 자신의 공간 정보를 저장하고 있어, Spatial Adjacency Matrix 가 제공되고 있음을 확인할 수 있다.

```
<mtci:CI>
  <mtci:view id="View1" style="position:absolute; width:3000px; height:780px" mtci:begin="0" mtci:end="indefinite">
    <mtci:divocation id="div1" style="position:absolute; left:0px; top:0px; width:1487px; height:855px" mtci:begin="0"
      mtci:end="indefinite" mtci:refDiv="Group1">
      <mtci:divocation id="div1.1" style="position:absolute; left:1085px; top:61px; width:1913px; height:719px" mtci:begin="0"
        mtci:end="indefinite" mtci:refDiv="Group2">
        <mtci:view id="View2" style="position:absolute; width:1487px; height:855px" mtci:begin="0" mtci:end="indefinite">
          <mtci:divocation id="div1.1.1" style="position:absolute; left:8px; top:0px; width:816px; height:612px" mtci:begin="0"
            mtci:end="indefinite" mtci:refDiv="Group2">
            <mtci:MediaSync refId="Media1" mtci:begin="0"/>
            <mtci:MediaSync refId="Media2" mtci:begin="0"/>
            <mtci:MediaSync refId="Media3" mtci:begin="0"/>
            <mtci:MediaSync refId="Media4" mtci:begin="0"/>
            <mtci:MediaSync refId="Media5" mtci:begin="0"/>
            <mtci:MediaSync refId="Media6" mtci:begin="0"/>
            <mtci:MediaSync refId="Media7" mtci:begin="0"/>
          </mtci:divocation>
        </mtci:view>
      </mtci:divocation>
    </mtci:divocation>
  </mtci:CI>
```

그림 8. CI 문서를 이용한 제안 기술 구현

HTML5 문서의 2 가지 'Group'을 이용하면 총 3 가지의 <view>를 구성하는 것이 가능하다. <그림 8>을 통해 CI 문서가 3 가지 <view>를 모두 전달하고 있는 것을 확인할 수 있다. 각각의 <view>는 제공되는 미디어의 전체 크기에 따라 다른 <width>, <height>를 가지며, 하위 단에는 'Group'을 <refDiv>를 통해 보유하고 있다.

서버가 MMT CI 를 통해 Spatial Adjacency Matrix 와 미디어 'Group' 정보를 전달할 수 있음을 구현하여 검증하였다. 그렇다면 사용자는 서버에게 전달받은 정보를 이용하여 원하는 관심 영역을 ROI message 로 요구할 수 있어야한다.

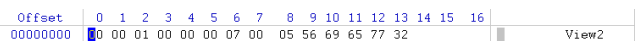


그림 9. ROI message 기반 <view> 요구 예시

먼저, 관심 영역이 서버가 제공한 <view> 중에서 존재할 경우를 생각할 수 있다. <그림 9>의 ROI message 는 7, 8 번째 Byte 에 존재하는 <view\_id\_length>의 값이 '5'로 <view>를 요구하는 ROI message 임을 알 수 있고, 제공된 <view> 중 'View2'를 요구하고 있다. 이때 <view\_id\_length>의 값이 0 을 갖는 경우는 <view>가 아닌 media 를 요구하는 것이다.

Offset	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16			
00000000	00	00	01	00	00	00	24	00	00	00	04	00	06	69	6D	61	67		\$	imag
00000017	65	33	00	06	69	6D	61	67	65	34	00	06	69	6D	61	67	65	e3	image4	image
00000034	35	00	06	69	6D	61	67	65	36									5	image6	

그림 10. ROI message 기반 미디어 집합 요구 예시

7, 8 번째 Byte 에 존재하는 <view\_id\_length>의 값에 '0'을 갖는 ROI message 의 예시가 <그림 10>에 나타나있다. 이는 ROI message 가 미디어들을 요구하고 있음을 뜻하며, 9, 10 번째 Byte 에 존재하는 <number of media> 값이 '4'이므로 총 4 개의 미디어를 요구하고 있음을 확인할 수 있다.

앞서 MMT CI 를 통해 Spatial adjacency Matrix 를 전송하고, 사용자가 이를 기반으로 ROI message 를 구성하는 것을 구현을 통해 검증하였다. 하지만 최종적으로는 실제 미디어를 이용하여 구현한 MMT CI 에 기술된 <view>들을 검증하는 것이 필요하다.



그림 11. 서버에서 제공한 <view> 중 'View1'의 검증.

<그림 11>은 실제 미디어를 이용해 MMT CI 에 기술된 <view>중 'View1'을 실험한 결과이다. 결과를 통해서 View1'에서는 2 개의 'Group'을 이용해 7 개의 미디어를 모두 제공하고 있는 점을 확인할 수 있다.



그림 12. 서버에서 제공한 <view> 중 'View2'의 검증.

<그림 12>는 실제 미디어를 이용해 MMT CI 에 기술된 <view>중 'View2'를 실험한 결과이다. 'View2'는 'Group1'을 제공하고 있는 <view>로, 'Group1'은 3 개의 미디어로 구성되어 있다. 이를 <그림 12>를 통해 확인할 수 있다.



그림 13. 서버에서 제공한 <view> 중 'View3'의 검증.

마지막으로 'View3'을 검증한다. 'View3'의 경우 'Group2'로 구성되어 있으며, 4 개의 미디어로 구성됨을 <그림 12>를 통해 확인할 수 있다. 상기 검증을 통하여 MMT CI 에 구성하고 있는 Media 들의 Spatial Adjacency Matrix 가 정확하게 기술되어 있음을 육안으로 확인할 수 있다. 또한 <view>의 크기가 미디어들의 전체 크기에 따라 정확히 기술되어 있음을 확인할 수 있다. 상기한 내용을 통해 최종적으로 제안한 기술들의 효용성을 검증하였다.

## 5. 결론

MMT CI 기반 Spatial Adjacency Matrix 전송 기법을 통해 사용자는 UWV 를 구성하는 미디어들의 공간적, 시간적 정보를 획득할 수 있고, 관심 영역의 미디어들을 전송 받는 것이 가능하다. 본 논문에서는 UWV 를 구성하는 미디어들을 'Group' 으로 명명하는 새로운 개념을 도입하여 Spatial Adjacency Matrix 를 전달하는 것을 제안하여 사용자에게 공간적, 시간적 정보를 전달했다. 또한 ROI Message 정의하여 사용자의 관심 영역을 서버 측에 알릴 수 있도록 제안하였다. 이는 한가지 UWV 미디어 집합을 이용하여 다양한 미디어를 소비할 수 있도록 만들 것이며, 이러한 변화는 사용자 친화적인 서비스에 한걸음 다가가 새로운 수익모델을 창출 할 수 있을 것이다. 마지막으로 추후 새롭게 정의한 ROI Message 를 통하여 서버에서 사용자에게 관심 영역의 미디어들을 제공하는 기법에 대한 연구를 진행할 계획이다.

\* 본 논문은 미래창조과학부 및 정보통신기술진흥센터의 정보통신·방송 연구개발사업의 일환으로 수행하였음. [2015-0-00231, 퍼즐형 Ultra-wide viewing 공간 미디어 생성 및 소비 기술 개발]

## 참고문헌

[1] TTA, Technical Report for High Quality Panorama Video, 2014

[2] Euisang Lee, Dongjin Kang, Jeonjo Kang and Kyuheon Kim, "GENERATION OF SPATIAL ADJACENCY MATRIX FOR ZIG-SAW LIKE ULTRA WIDE VIEWING DATA", EEECS 2017.

[3] ISO/IEC JTC 1/SC 29 23008-1:2014, Informaiton technology - High efficiency coding and media delivery in heterogeneous environments - Part 1 : MPEG Media Transport(MMT)

[4] ISO/IEC JTC 1/SC 29 23008-11:2014, Informaiton technology - High efficiency coding and media delivery in heterogeneous environments - Part 11 : Composition Information