

## MMT-CI 기반 멀티 소스를 활용한 효율적 장면구성정보 설계

홍 지 예, 김 규 현

경희대학교

chcojiye@khu.ac.kr, kyuheonkim@khu.ac.kr

Design the Scene Coposition-Information using Multi-Source  
based on MMT-CI

Hong, Ji-Ye Kim, Kyuheon

Kyung-Hee University

## 요 약

방송융합시대가 도래함에 따라 방송뿐만 아니라 IP망을 통해 콘텐츠를 소비할 수 있는 스마트TV의 보급이 급속히 확산되고 있다. 또한 영상과 음악 및 광고와 같은 멀티미디어 콘텐츠의 소비 환경이 한 화면에 한 가지의 미디어 서비스만 제공받는 환경에서 다양한 정보를 동시에 소비할 수 있는 환경으로 확대되면서 다양한 정보를 하나의 화면에서 혼용적으로 전달하기 위한 장면구성정보가 필요하게 되었다. 장면구성정보는 미디어가 소비되는 특정 시간과 해당 공간의 정보를 기술하므로써 다수의 미디어를 복합적으로 소비하는 방법으로, IP 기반에서도 여러 개의 화면을 통해 여러 소스의 콘텐츠를 소비하여 다양한 사용자 경험(User Experience)을 제공하고자 ISO/IEC JTC1/SC29/WG11(별칭 MPEG)에서는 방송서비스에서 방송망 및 IP망을 활용하여 멀티미디어 데이터를 전송할 때 사용하는 MPEG Media Transport(MMT)를 기반으로 장면구성정보를 제공하기 위한 MMT-Composition Information(MMT-CI)를 기술표준화하였다. 본 논문에서는 MMT-CI의 표준문서에 따라 W3C의 웹언어인 HTML5와 확장성 언어인 XML을 이용하여 멀티소스를 활용한 장면구성정보를 기술하는 것을 제안하고 장면구성서비스에 적용을 위한 플레이어를 구현 및 검증한다.

## 1. 서 론

기존의 멀티미디어 콘텐츠의 소비환경은 TV를 통해 단방향으로 제공받거나 인터넷을 통해 파일기반의 다운로드를 받는 방식이었다. 하지만 최근 광대역의 초고속 통신망이 보급됨에 따라 콘텐츠 소비환경은 IP망과 방송망을 이용한 양방향서비스와 인터넷을 통한 고화질 스트리밍 서비스로 변모하게 되었다. 이에 방송 이외의 매체를 통한 비디오 콘텐츠 소비가 증가하게 되었고 IPTV와 N스크린 서비스의 활성화로 전통적인 방송 산업구조에도 변화가 일어나면서 IP기반의 콘텐츠 소비환경과 방송융합의 시대가 가속화단계에 접어들었다<sup>[1]</sup>. IP기반의 콘텐츠 소비환경에 맞춰 ISO/IEC의 멀티미디어 국제표준화 기구 MPEG(Moving Picture Expert Group)에서는 차세대 방송시스템 표준으로 MPEG Media Transport(MMT)를 기술표준화 하였다<sup>[2]</sup>.

이와 더불어 단말기와 통신 기술의 발달로 최근 급격한 모바일기기의 확산됨에 따라 콘텐츠 소비환경이 TV에서 다양한 멀티미디어로 변화하고 있다. 한 가지 미디어를 단방향으로 제공해주는 환경이 아니라 다양한 정보를 동시에 소비할 수 있는 환경으로 발전하였고, 하나 이

상의 콘텐츠를 동시에 소비하기 위해 화면을 구성하여 다양한 콘텐츠와 각종 광고 및 커머스 서비스를 제공하고자 한다. 이와 같은 변화된 환경에 적합한 기술을 제공하기 위하여, IP기반의 환경에서 소비자의 요구에 따라 멀티미디어 콘텐츠의 다양한 장면구성을 위해 ISO/IEC 23008-11 MPEG-H(Moving Picture Expert Group - H) Part 11의 MMT-Composition Information에서는 장면구성정보 기술에 관한 기술규격을 표준화하고 있다<sup>[3]</sup>.

본 논문에서는 MMT 기반의 CI 문서를 통해 장면구성정보를 제공하여 장면을 구성하는 다수의 미디어 객체들을 분류하고 이에 대한 시공간적 정보와 상호간의 연관성, 사용자의 요청에 따른 이벤트 정보를 처리하는 방법을 제안하고 플레이어를 구현하여 이를 검증해냈다.

이와 같은 내용을 기반으로 본 논문의 구성은 다음과 같다. 2절에서는 장면구성정보에서 사용하는 웹문서 세부 기술에 대해 소개한다. 3절에서는 MMT기반의 장면기술정보를 세부적으로 설계할 것이다. 4절에서는 본 논문에서 제안하는 CI를 이용하여 장면을 기술하는 플레이어를 구현하여 검증하고, 마지막으로 5절에서는 본 논문에 대한 결론을 맺는다.

## 2. 장면구성정보를 위한 웹문서 세부 기술 소개

HTML5문서와 XML문서를 통해 장면구성정보 기술을 설계하기 위해, W3C의 HTML5, XML과 DOM 에 대해 소개한다.

### 가. HTML5

HTML5(Hyper Text Markup Language 5)는 웹문서를 만들기 위한 마크업 언어의 한 종류로 W3C에서 제정한 HTML의 5번째 버전이다. 현재는 Proposed Recommendation 16 September 2014까지 공표된 상태이다<sup>[4]</sup>. HTML5는 단일 기술이 아니라 다양하게 명세된 기술들의 표준화된 규격으로 자바스크립트 API확장을 정의하여 과거 외부 플러그인 확장들을 통해서만 제공이 가능하던 오디오, 비디오 처리의 기능을 제약 없이 웹 브라우저가 표준화된 방식으로 제공할 수 있도록 하는 특징이 있다.

HTML5는 웹페이지에서 Flash나 Javascript없이 다양한 미디어 멀티 소스의 공간구성이 가능하게 하나 구현하기 복잡하고 화면 전환 시 시간적 구성은 불가능하다는 단점이 있다. MMT-CI에서 장면구성정보에서 HTML5는 공간정보구성에만 이용되고 다음 절에서 소개하는 XML을 통해 HTML5의 시간구성의 한계점을 극복할 것이다.

### 나. XML

XML(Extensible Markup Language)은 웹에서 사용하는 확장성 마크업 언어이다. HTML5가 웹 문서 제작을 위한 마크업 언어이기 때문에 멀티소스간의 시간구성 방법이 정의되어 있지 않으며, 스크립트 및 스타일 언어 사용으로 인한 복잡성을 가진다. XML은 문서를 사람과 기계 모두가 읽을 수 있는 형식으로 부호화하기 때문에 수많은 종류의 데이터를 기술하는 데에 적합하다<sup>[5]</sup>. 이와 같은 문서 특성으로 인해 장면구성정보에서 시간정보를 기술하기 위해 CI문서에 도입하게 되었다.

### 다. DOM(Document Object Model)

HTML5, XML 문서는 텍스트 형태의 논리적 구조로 나타냄으로써 사람이 읽을 수 있으나 컴퓨터는 해당 문서의 텍스트 정보만으로 구조적인 내용에 대해 인식이 불가능하다. 이에 W3C에서는 DOM(Document Object Model)이라는 프로그래밍 인터페이스 문서를 권고함으로써, 문서의 각 요소들을 구조적인 객체 표현방식인 트리구조로 접근이 가능 하게 하는 API를 제공한다<sup>[6]</sup>.

따라서, DOM을 이용하여 HTML, XML문서의 요소들과 속성들은 각각의 노드를 형성하여 부모 자식의 계층 형태로 메모리에 저장할 수 있다.

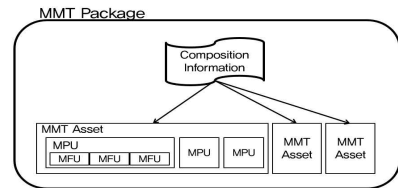
## 3. MMT-CI 기반 장면구성정보

MMT기반의 장면구성정보 기술을 설계하기 위해, MMT의 데이터모델과 MMT-CI의 문서구조, 장면구성에 관한 개념을 설명한다.

### 가. MMT 데이터 모델

MMT(MPEG Media Transport)는 IP기반의 미디어 전송 서비스를 가능하게 하는 MPEG의 차세대 방송 시스템 표준이다. MMT를 사용하여 전송되는 미디어 데이터의 논리적 집합을 Package라고 하며 방송환경에서는 프로그램을 하나의 Package로 정의할 수 있다.

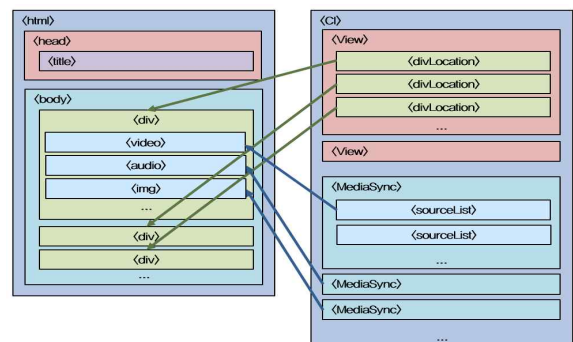
Package에서는 비디오, 오디오, 데이터를 포함한 다양한 멀티 소스들을 각각의 Asset으로 정의한다. <그림1>에서 보는 바와 같이 MMT Package는 MMT Asset의 집합으로 구성되어 있다. 각 Asset은 전송 및 재생/저장의 최소단위인 MPU(Media Processing Unit)들의 집합으로 구성되고 MPU는 frame단위의 MFU(Media Fragment Unit)들로 이루어져 있다. <그림1>에서와 같이 한 Package는 CI문서를 반드시 하나 이상 포함하고 CI문서는 Package안의 모든 멀티미디어 소스인 Asset간의 관계와 MPU의 시간 및 공간관계에 대한 정보를 가지고 있다.



<그림1. MMT 데이터 모델 중 CI의 역할>

### 나. MMT-CI의 문서구조

CI문서는 HTML5와 XML 신텍스에 따라 작성되어 각각 2개의 문서로 구성된다. HTML5의 공간구성 기능과 XML의 시간구성 기능을 복합적으로 사용하여 MMT의 asset들이나 MPU간의 시간, 공간 정보를 포함한다. <그림2>은 HTML5문서와 CI문서의 구성요소와 관계를 나타내고 있다.



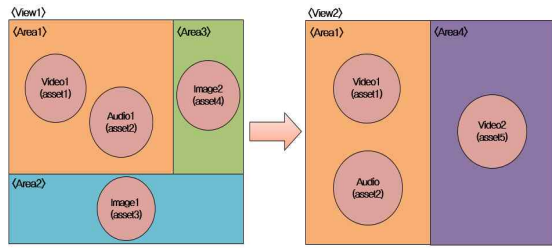
<그림2. MMT-CI의 HTML, CI 문서구조와 상호관계>

<그림2>의 좌측 HTML문서의 <body>에는 Asset의 성격을 정의한다. HTML5의 화면을 분할하는 일반적인 요소(element)인 <div>의 하위요소에 멀티 소스인 <video>, <audio>, <image>등의 요소를 정의한 후 미디어 요소의 공간정보를 속성(attribute)에 기술한다. XML Schema로 작성된 <그림2> 우측의 CI문서에는 하나의 장면을 나타내는 <View>요소를 가장 먼저 기술하며 하위항목 요소로는 <divLocation>요소가 정의되어 있다. 이 요소는 HTML5 문서에서 하나의 MMT Asset 혹은 하나의 MPU에 해당하는 멀티소스 요소들의 데이터 값을 참조하여 화면에 렌더링하는 역할을 한다. 다음은 <MediaSync>요소로 HTML5에 미디어 요소들의 시간정보를 XML 데이터인 CI문서를 통해 참조하여 해당 소스의 시작시간과 종료시간을 기술한다.

### 다. MMT-CI에 의한 장면구성

장면이란 단말의 화면에서 표현할 수 있는 비디오, 오디오, 이미지, 텍스트 등과 같은 멀티미디어들의 구조적 형태를 뜻

하며 이는 장면 구성 정보를 통해 각 객체들 간의 시공간적 소비 방안을 제시한다.



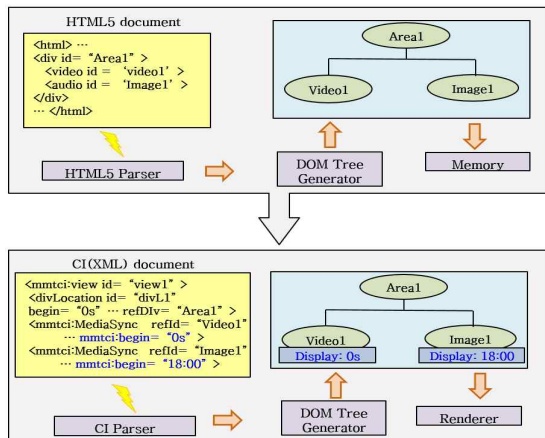
<그림3. 장면구성 개념도>

본 논문에서는 HTML5가 제공하는 공간구성을 활용하면서 HTML5문서와 CI문서 하나로 다수의 장면을 기술할 수 있도록 <그림2>의 요소들을 기반으로 본 논문에서는 다음의 <그림3>와 같은 장면구성 개념을 도입한다.

하나의 장면은 View개념으로 대체되며, View는 Area의 조합으로 이루어진다. Area는 장면의 일부를 독립적으로 구성하며 View가 변화할 때 동일한 Area를 사용가능하다. 미디어는 Area내에서 보여 지게 되며, 이러한 구조는 반복적으로 나타나는 미디어에 대한 중복을 줄이고, 장면의 일부를 독립적으로 변경하는 것에 용이하다.

#### 4. 장면구성시스템 제안 및 플레이어구현

##### 가. 장면구성시스템 설계

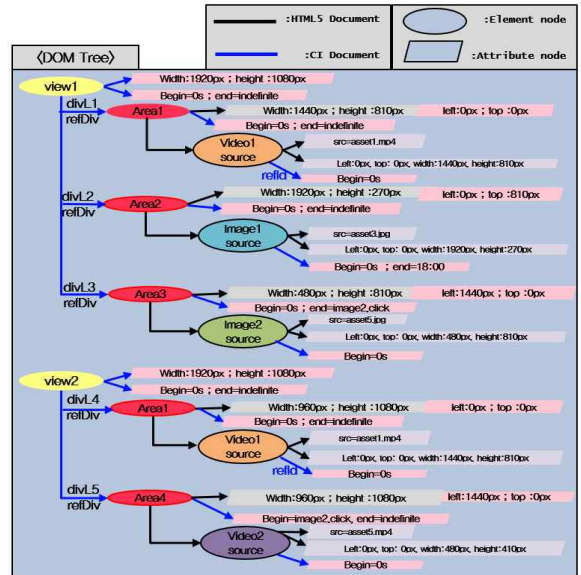


<그림4. CI Processing Engine에서 장면구성 시스템 구성 예시>

장면을 구성하기 위해 CI문서를 처리하기 위한 방법으로서 <그림4>에서와 같이 장면구성정보 처리 엔진(CI Processing Engine)을 제안한다. HTML5 문서는 비디오와 오디오, 이미지 등 각종 멀티소스의 공간 초기정보를 담고 있기 때문에 가장 먼저 Parser(분석기)를 통해 분석된다. HTML의 각 요소마다 DOM Tree의 node를 생성하여 계층적 트리구조로 구성한 후 메모리에 적재한다. <그림4>에서는 'Areal' 과 'Video1', 'Image1' 이 계층적 노드로 DOM Tree를 구성하였다.

다음 CI문서가 CI Processing Engine에서 CI Parser를 통해 분석된다. CI문서에서 분석한 'refDiv' 속성값과 기존 메모리상의 'div' 속성값을 매칭하여 해당 멀티소스의 시간

정보 값을 각 DOM Tree Node에 업데이트시킨다. <그림5>은 파싱된 정보를 바탕으로 DOM tree구조로 메모리에 저장되어 있는 객체들을 업데이트 시키는 과정을 보여준다.



<그림5. DOM Tree 구성 예시>

<그림5>에서 HTML5 문서에 의해 생성된 요소와 속성들은 흑색 선으로 표시했다. CI문서를 통해 업데이트 된 요소와 속성값들은 청색 선으로 나타내었다. 이 때 주로 업데이트 되는 내용은 미디어 데이터가 재생되는 시간이다. 다음과 같이 생성된 계층적으로 구성된 DOM-tree는 메모리 공간을 효율적으로 사용할 수 있게 하여 화면구성시간을 단축시킨다. 구성된 DOM-tree는 Renderer로 전달되며 View의 변화에 따라 미디어 재생 위치 및 시점에 맞춰 미디어 소비가 이루어지게 된다.

##### 나. MMT-CI Player 구현 및 검증

장면구성 정보를 담고 있는 HTML5, CI 문서를 재생하는 CI 플레이어를 구성하였다. CI플레이어에서는 구현을 목적으로 미디어는 비디오만, 시간구성 정보는 상대시간의 초 단위만 사용가능 하도록 하였다.



<그림6. MMT CI Player의 사용자 인터페이스>

먼저 ① 버튼으로 HTML5을 입력받아 HTML5 파서로 분석하여 장면구성정보를 DOM-tree Generator로 전달한다. Generator는 DOM Tree를 생성하여 메모리에 저장한다. 다음 ② 버튼으로 update할 CI문서를 불러온다. CI 문서도 CI 파서(Parser)를 통해 분석되고 DOM-tree Generator에 전달되어 기존에 메모리에 있는 내용을 업데이트 시킨다. 업데이트 된 내용은 ③ Play버튼으로 비디오 디코더와 렌더러(renderer)에 전달되어 장면구성정보의 공간구성에 따라 알맞은 영역에 표시되도록 한다. 해당 실험을 위해 사용된 HTML5와 CI문서는 <그림7>과 같다.

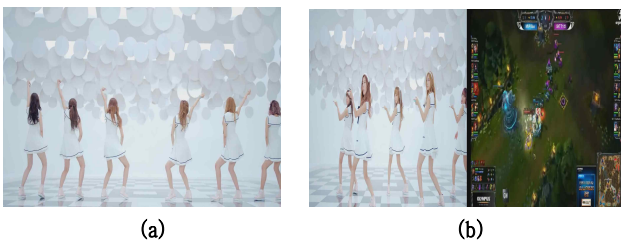
```

<!DOCTYPE html>
<html xmlns="http://www.w3.org/1999/xhtml">
<head>
<title>Test</title>
</head>
<body>
<div id="Area1" style="position:absolute; width:1920px; height:1080px">
<video id="video1" src="asset1.mp4" style="position:absolute; left:0px; top:0px; width:960px; height:1080px"/>
</div>
<div id="Area2" style="position:absolute; width:1920px; height:1080px">
<video id="video2" src="asset2.mp4" style="position:absolute; left:0px; top:0px; width:960px; height:1080px"/>
</div>
</body>
</html>

<mmtci:CI>
<mmtci:view id="View1" style="position:absolute; width:1920px; height:1080px" mmtci:begin="0s" mmtci:end="5s"> (ㄱ)
<mmtci:divLocation id="divL1" style="position:absolute; left:0px; top:0px; width:1920px; height:1080px" mmtci:begin="0s" mmtci:end="indefinite"
mmtci:refDiv="Area1"/>
</mmtci:view>
<mmtci:view id="View2" style="position:absolute; width:1920px; height:1080px" mmtci:begin="5s" mmtci:end="10s"> (ㄴ)
<mmtci:divLocation id="divL2" style="position:absolute; left:0px; top:0px; width:960px; height:1080px" mmtci:begin="0s" mmtci:end="indefinite"
mmtci:refDiv="Area1"/>
<mmtci:divLocation id="divL3" style="position:absolute; left:960px; top:0px; width:960px; height:1080px" mmtci:begin="0s" mmtci:end="indefinite"
mmtci:refDiv="Area2"/>
</mmtci:view>
<mmtci:MediaSync refId="Video1" mmtci:begin="0s"/>
<mmtci:MediaSync refId="Video2" mmtci:begin="Video1.begin"/> (ㄷ)
</mmtci:CI>
    
```

〈그림7. 실험용 HTML5 및 CI 문서〉

재생시작부터 5초간 (ㄱ)의 “view1”에 의해 “divL1”이 활성화되어 “Area1”의 “asset1.mp4”가 1920\*1080의 크기로 전체화면에 재생된다. 다음 5초 이후에는 (ㄴ)의 “view2”로 장면이 전환되며 “divL2” “divL3”이 활성화된다. 이 때, “divL2”는 “Area1”을 참조하므로 “asset1.mp4”는 끊김 없이 재생되고 화면의 크기만 960\*1080으로 변하게 된다. (ㄷ)은 “divL3”에 참조되어 있는 “Area2”의 “asset2.mp4”를 재생하게 된다. 이 때 video1과 video2의 동기를 맞추기 위해 (ㄷ)의 MediaSync요소를 통해 video2의 시작시간을 video1의 시작시간으로 설정하였다.



〈그림8. (a)시작 후 5초간 영상, (b)5초 후 화면 분할 영상〉

이와 같은 HTML5 및 CI 문서 기반의 장면구성시스템이 본 논문에서 제안한 장면구성 방법에 따라 <그림8>의 문서들을 통해 미디어를 재생하는 장면을 나타냈다. <그림8>의 (a)에서 “View1”의 “Area1”에 있는 “asset1.mp4”를 재생하는 것을 확인할 수 있었고, <그림8>의 (b)에서 “View1”이 5초에 종료되고 “View2”에 의해 “Area1”의 “asset1.mp4”와 “Area2”의 asset2.mp4가 공간구성에 따라 재생되는 것을 확인할 수 있었으며, “asset1.mp4”가 “View1”에서 “View2”로 전환되었을 때 멈추거나 처음으로 돌아가는 현상 없이 재생되는 것을 확인하였다.

## 5. 결론

MMT기반의 장면구성기술 제안을 통해 복수 개의 미디어를 MMT Asset단위로 소비할 때 미디어의 공간적, 시간적 구성정보를 제공함으로써 IP나 하이브리드 방송과 같이 다수의 미디어 전송 환경에서 사용이 가능하다.

본 논문에서는 HTML5 문서와 CI 문서로 비디오, 오디오, 비디오, 오디오, 이미지, 각종 광고배너 등과 같은 멀티소스들에 대한 장면 구성 기술을 DOM을 통해 계층적인 형태로 기술함으로써 독립된 객체간의 상호관계를 정립하였고, MMT Asset들을 재생할 수 있는 플레이어를 구현 및 검증하였다. 또한 객체 간의 시간적, 공간적인 배치와 사용자의 입력이나 시간에 따른 객체의 동작의 접근 및 조작이 가능한 기술을 제공할 수 있다. 이를 통해 차세대 방송시스템인 MMT의 도입과 함께 MMT기반의 방송에 대한 다양한 장면구성에 적용이 가능하고 더불어 하이브리드 방송환경에서 멀티뷰, 서라운드 뷰와 같은 서비스의 수익모델을 창출할 수 있을 것이다. 나아가 시청자의 기호에 맞는 멀티미디어 콘텐츠들을 선택할 수 있는 리치미디어 서비스를 가능하게 할 것이다.

## 6. 참고문헌

- [1] KOCCA, ‘멀티스크린 시대로의 전환과 콘텐츠 이용행태의 변화’, 한국콘텐츠진흥원, 2014, 8월
- [2] 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)
- [3] ISO/IEC JTC 1/SC 29 23008-11:2014, Informaiton technology - High efficiency coding and media delivery in heterogeneous environments - Part 11 : Composition Information
- [4] W3C, “HyperText Markup Language (HTML5) Proposed Recommendation, September 2014
- [5] W3C, “EXtensible Markup Language (XML) 1.0 (Fifth Edition) W3C Recommendation, November 2008
- [6] W3C, “Document Object Model (DOM) The WebApps WG Drives DOM Specifications. January 6th, 2009