

# 다양한 단말에 맞춤형 서비스를 위한 적응적 장면 기술 방법 연구

김병철 김규헌

경희대학교

coel@khu.ac.kr kyuheon.kim@googlemail.com

## Adaptive Scene Description Method for Heterogeneous Devices

Byungchul Kim Kyuheon Kim

KyungHee University

### 요약

현재의 멀티미디어 서비스는 기존의 텍스트 위주의 사용자 환경에서 벗어나 오디오, 비디오뿐만 아닌 이미지, 그래픽, 텍스트 및 미디어 관련 정보 등의 다양한 데이터를 포함하는 멀티미디어 서비스 환경으로 변화하고 있다. 이러한 멀티미디어 서비스는 과거 TV와 컴퓨터 등 한정된 기기에서만 제공되었지만, 최근 휴대폰, MP3 플레이어, PMP 등 미디어 기기의 발전에 따라, 다양한 단말을 대상으로 서비스를 확장하고 있다.

이러한 멀티미디어 서비스를 다양한 단말을 대상으로 제공하기 위하여, 기존의 기술에서는 동일한 콘텐츠를 각각의 단말의 환경에 따라 서로 다른 서비스를 통해 제공하였다. 그러나 IPTV와 홈 네트워크 시스템 등 방송과 통신이 융합되고, 유선과 무선 간에 자유롭게 데이터를 주고받는 환경에서 단말마다 서로 다른 서비스를 제공하는 것은 서비스 시나리오의 복잡도를 높이고, 다양한 단말을 사용하는 사용자에게 통합되지 않은 서비스를 제공 받게 되는 결과를 가져오게 된다.

이에 본 논문에서는 콘텐츠의 메타데이터를 활용하여 동일한 멀티미디어 서비스를 단말기의 다양한 환경에 따라 맞춤형 서비스를 제공할 수 있는 방법인 적응적 장면 기술을 제시하고, 그 효율성에 관하여 검증한다.

### 1. 서론

최근 멀티미디어 기술은 컴퓨터와 미디어 서비스의 급속한 발전으로 인해 기존의 텍스트 위주의 사용자 환경에서 벗어나 오디오 및 비디오뿐만 아닌 이미지, 그래픽, 텍스트 및 미디어 관련 정보 등의 다양한 데이터를 포함하는 멀티미디어 서비스 환경으로 변화하고 있다.

또한 과거 TV와 컴퓨터로만 서비스 될 수 있었던 멀티미디어는, 최근 휴대폰, MP3 플레이어, PMP 등 전자기기의 발전에 따라, 다양한 단말에서 멀티미디어 서비스를 제공 하는 기술로 발전되고 있다.

현재 여러 종류의 휴대용, 가정용 멀티미디어 기기가 시장에 선보이고 있으며, 이러한 단말기들은 디스플레이 크기, CPU, 메모리, 배터리 등의 성능 요건이 서로 다르다. 그리하여, 동일한 콘텐츠를 재생하는 경우 단말기의 성능에 따라 제공받은 콘텐츠의 서비스 상황이 각각 다르다. 즉 TV와 같이 방송이 주가 되는 단말에서는 HD화질과 같은 고해상도의 서비스를, 휴대폰과 같이 방송이 보조적인 성격을 보이는 단말에서는 저 해상도 서비스를 제공하는 것처럼, 단말의 성능과 목적에 따라 각각 다른 서비스를 해 주어야 된다. 하지만 각기 다른 단말기의 성능에

대한 정보가 없는 경우, One source, Multi use의 시나리오 상에서는 위와 같은 형태의 서비스 제공과 멀티미디어 콘텐츠 재생이 불가능하다.

따라서 이러한 문제를 해결하기 위해 하나의 콘텐츠를 다양한 단말기를 대상으로 서비스하고자 할 때, 하나의 멀티미디어 콘텐츠를 단말기의 성능 요건에 맞추어 서비스 할 수 있는 방법을 필요로 하고 있기에, 본 논문에서는 장면 구성 기술을 통해 동일한 멀티미디어 서비스로 단말기의 다양한 환경에 따라 맞춤형 서비스를 제공할 수 있는 방법에 대해서 제시하고, 그 효율성에 관하여 논하고자 한다.

### 2. 본론

#### 1. LASeR 기술개발 배경 및 목적

정보통신 기술의 발전에 따라 사용자는 이제 단순한 AV 위주의 멀티미디어 서비스에서 벗어나 혼합된 여러 가지의 미디어가 사용자의 요구사항을 받아들이는 다양한 대화형 서비스를 제공하는 형식으로 발전하고 있다.[1] 이러한 기술을 리치미디어 (Richmedia) 서비스라

하며 서비스 기술 개발이 발전하고 확장됨에 따라 통신과 방송 영역에서 활용될 기술 개발이 필요하게 되었다.

이러한 리치미디어 서비스는 기존의 MPEG-4에서 제공하는 장면 구성 기술인 Binary Format for Scene (BIFS)를 통해 가능하였다.[1] 지상파 DMB에서 TV프로그램과 데이터서비스를 결합시키는 BIFS는 여러 개의 객체(영상화면, 텍스트 화면 등)에 대해 각 객체가 어디에 위치할지 언제 등장할 지를 설정해 두는 방식이다. 즉 TV화면에 다른 데이터를 얻을 수 있고, 이 데이터를 선택하면 다른 정보 화면이 뜨게 하는 식으로 연동형 데이터서비스를 가능케 하는 기술이다. 그러나 BIFS의 경우, 해당 기술이 Digital Television(DTV)의 Set Top Box(STB) 또는 인터넷 방송을 그 서비스 대상으로 여기고 있었기에 그 복잡도가 상당히 높아 연산능력이 충분한 단말에 적용된 기술로 메모리 용량이나 저장 용량과 성능이 부족한 휴대폰 같은 모바일 기기에서는 한계가 있다.

이를 보완하기 위해 MPEG은 풍부한 화면 구성과 압축의 효율성 두 가지를 모두 고려한 새로운 규격을 만들었고 이것이 바로 Lightweight Application Scene Representation (LASeR)이다.[2] LASeR는 2004년부터 본격적인 표준화 작업이 진행되었으며, 2006년 1월 첫 번째 버전의 표준 작업이 완료되었다. 또한 2008년과 2009년에는 멀티미디어가 다양한 종류의 단말기에서 소비되는 환경을 고려하여, 하나의 콘텐츠를 다양한 단말에 맞춤형 서비스를 제공하는 Scene Adaptation 기능을 추가하였고, 2009년에는 장면서술정보에서 구조화된 정보의 일부분을 참조하여 장면을 효과적으로 표현해주는 PMSI 기술을 LASeR에 추가하는 작업을 진행하여 표준 문서 작업은 완료하였다. 또한 현재 2010년에는 위의 PMSI 기술을 접목시킨 LASeR 기능을 점검하는 Reference Software 작업이 진행되고 있다.

## 2. 장면 기술 표준 LASeR

LASeR는 멀티미디어 서비스를 위한 장면 기술 정보를 바이너리 형식으로 제공하는 방법을 정의한 규격으로써, <그림 1>에서 나타나는 바와 같이 크게 멀티미디어 장면 기술 부분과 미디어로 구성되어 있다. LASeR를 구성하는 미디어로는 정지 영상, 동영상, 오디오 및 폰트를 지원하며, 장면 기술은 텍스트, 그래픽스, 애니메이션과 각각의 객체인 미디어에 대한 시공간적 구성 정보와 대화형 기능 등으로 구성되어 있다. [2]

이러한 장면 구성 정보는 즉 LASeR 장면 기술은 크게 아래와 같이 3가지 기능제공을 목적한다.

- 1) 미디어 및 그래픽스와 같은 장면을 구성하는 요소들의 시공간적 구성 방법
  - 2) 하나의 장면을 구성하는 요소 간의 상호 작용 방법
  - 3) 장면이 변경될 경우, 변경 부분을 반영하는 방법
- 상기의 기능을 제공하기 위하여 LASeR에서는 World

Wide Web Consortium (W3C)의 주도하에 개발된 오픈 표준의 벡터 그래픽 파일 형식인 Scalable Vector Graphics (SVG)에 기반을 두어 장면을 구성하며, SVG 규격에서 지원하지 않은 기능들은 LASeR extensions를 통해 지원하도록 하였다. 초기에 LASeR는 SVG 1.1 버전을 기반으로 하였지만, SVG profile에 비해 양은 작지만 device GUI, 2D Game, 대기화면 서비스 등과 같은 Visual Application 적용을 위한 리치미디어와 강력한 Interactivity를 지원하는 모바일 버전의 SVG Tiny 1.2버전을 수용하였다.[3]

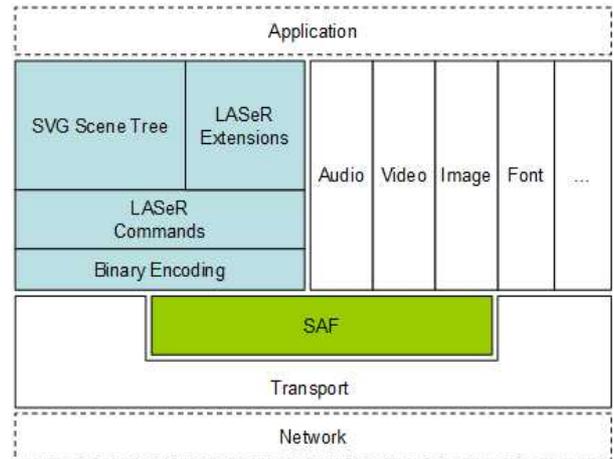


그림 1 LASeR와 SAF 구조

또한 SVG는 XML로 기술하므로 웹 브라우저 상에서 열람할 수 있고, 문서 편집기 등에서 편집을 할 수 있는 특징이 있으며, 하이퍼링크를 이미지에 걸거나, 자바스크립트 등과 연동시킬 수 있는 기능 등을 포함하고, 벡터 그래픽스이므로 확대나 축소를 해도 화질에 변화가 없는 장점을 가지고 있다.

이러한 SVG의 장면 구성요소는 <그림 2>에서 보는 바와 같이 도형, 텍스트, 이미지, 등의 다양한 조합과 애니메이션, 하이퍼링크, interactivity, Scripting등의 기능을 XML기반의 파일 형식으로 표현한 것이다.[4]

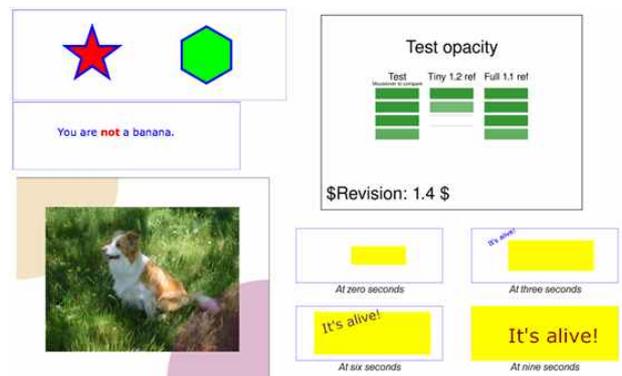


그림 2 SVG 장면 구성 요소

따라서 LASeR extensions은 상기의 SVG의 장면 구성 요소에 주로 interactivity와 A/V 콘텐츠의 구성요소를 표현하기 위해 필요한 장면 구성 요소와 좀 더 다양

한 기능을 제공하기 위해 추가한 것으로 볼 수 있다.

LASeR는 위의 장면 구성 요소들을 MPEG-4 BIFS에서 활용중인 'command' 방식을 활용함으로써 현재 보이고 있는 장면에 나타난 미디어 객체를 삭제/변경할 수 있으며, 또한 새로운 미디어 객체를 삽입할 수 있게 함으로써 보다 효율적으로 현재 장면에 대한 변경을 가능하게 한다.

또한 LASeR는 'command' 형식으로 구성된 장면 요소를 스트리밍 서비스에 적합하게 이진화 하는 과정을 거쳐, 별도의 포맷 형식인 Simple Aggregation Format (SAF)이라는 전송 규격 형태로 정의하여 전송을 한다.

SAF 규격은 LASeR 콘텐츠를 장면 묘사와 각 구성 미디어 요소를 포함한 포괄적 데이터의 이진화된 비트스트림을 전송하기 위한 전송 규격으로써 모바일 환경과 같은 작은 전송 대역폭에 적합하게 전송하기 위하여 다중화에 필요한 헤더 정보를 최적화하여 제공한다.

마지막으로 LASeR 콘텐츠는 MPEG-4의 ISO Base Media File Format을 기반으로 하는 MP4 file 형태로 저장된다. 이 파일 포맷은 다른 여러 파일 포맷기술의 기본 형태이며 가장 일반적인 포맷 기술로 각각의 객체를 track별로 기술하여 저장하기에, 위의 장면기술은 video, audio, text 등과 같이 각각의 media와는 별개로 scene track에 저장된다. 이 장면기술을 포함하는 track은 미디어의 타입을 나타내는 handler\_type이 "sdsm"이라는 장면기술타입 (Scene description type)의 값을 가지고, LASeR Header정보는 각 트랙의 헤더 정보를 기술하는 ConfigurationBox에 기술함으로써 이진화된 LASeR 스트림을 파일 포맷 형태로 저장 할 수 있다.

### 3. 적응적 장면 기술

본 논문에서 제안하는 기술은 콘텐츠 제공자가 서비스를 제공하고자 할 때, 하나의 콘텐츠와 함께 LASeR 장면 구성 정보에 단말의 성능 정보를 추가하여 서비스를 제공함으로써, 사용자는 단말기와 성능에 따라 서로 다른 서비스를 제공받을 수 있는 것을 목표로 하고 있다.

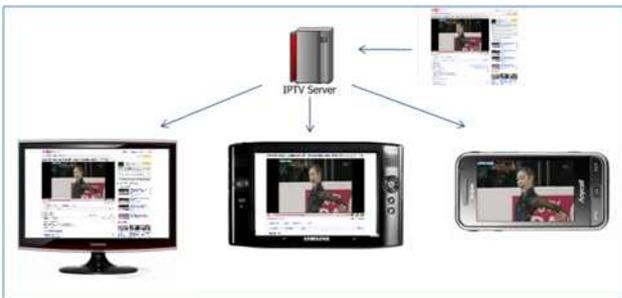


그림 3 적응적 장면 기술사용의 예

즉 <그림 3>에서 나타내는 바와 같이 같은 비디오, 이미지, 텍스트 등을 포함하는 리치미디어를 서비스 하더라도, 각 단말에서는 자신의 단말의 화면 크기와 성능에 맞춰 가능한 부분만을 재생하게 됨으로써, 각 단말에 따라

서로 다른 콘텐츠를 제작할 필요성이 줄어들고, 통합된 서비스를 제공하게 되어 서비스의 복잡도를 줄일 수 있다.

장면 기술 요소에 추가 되는 단말의 성능 정보는 Scene Adaptation이라 명기하며, 단말의 화면 크기를 나타내는 minDisplay size와 사용가능한 메모리 량을 나타내는 minMemory, 단말의 resolution 등이 있다. 이러한 성능 정보를 LASeR의 장면 구성 요소에 접목시켜, 기존의 LASeR 재생기에서 해당 기술을 추가하였으며, 이중 minDisplay size 속성을 통한 실험을 통해 Scene Adaptation을 이용한 서비스의 효율을 검증하고자 한다.

또한 Scene Adaptation 기술은 다양한 단말기를 대상으로 하여 제공 될 수 있기에, 더욱더 원활한 리치미디어 서비스의 주요 기술이 될 것이며, 콘텐츠 제공자의 콘텐츠 제작을 간편하게 하여 해당 미디어 서비스 산업에 파급 효과가 클 것으로 예상된다. 또한 현재 MPEG와의 다른 기관과 단체에서도 이와 관련한 표준화 활동이 활발히 진행되고 있어, 새로운 미디어 관련 기술로 주목을 받고 있다.

### 4. Scene Adaptation 실험 결과

본 실험은 단말에 맞는 크기의 video콘텐츠를 제공하여 얻을 수 있는 효과에 대해서 실험을 하였다.

이 실험은 표 1에서 보는 바와 같이 최소화 화면의 크기를 20인치 이상에서 8인치까지 2인치 간격으로 하여, 같은 비디오 콘텐츠를 재생하되 가로와 세로의 크기를 해당인치에 맞추어 재생하도록 하였다.

또한 실험 대상은 데스크톱에서 이루어 졌으며 LASeR 재생기에서 임의로 화면의 크기를 변경하여 실험을 수행하였다.

표 1 minDisplaySize에 따른 비디오 크기 조절의 예

```

<svg width="900" height="600" viewBox="0 0 900 600">
  <g minDisplaySize="20">
    <video id="video_sample1" x="0" y="0" width="800" height="600" xlink:href="video1"/>
  </g>
  <g minDisplaySize="18">
    <video id="video_sample1" x="0" y="0" width="750" height="550" xlink:href="video1"/>
  </g>
  <g minDisplaySize="16">
    <video id="video_sample1" x="0" y="0" width="700" height="500" xlink:href="video1"/>
  </g>
  <g minDisplaySize="14">
    <video id="video_sample1" x="0" y="0" width="650" height="450" xlink:href="video1"/>
  </g>
  <g minDisplaySize="12">
    <video id="video_sample1" x="0" y="0" width="600" height="400" xlink:href="video1"/>
  </g>
  <g minDisplaySize="10">
    <video id="video_sample1" x="0" y="0" width="550" height="350" xlink:href="video1"/>
  </g>
  <g minDisplaySize="8">
    <video id="video_sample1" x="0" y="0" width="500" height="300" xlink:href="video1"/>
  </g>
</svg>

```

LASeR 재생기에서 임의의 변경한 화면 크기는 21인치 부터 2인치 단위로 9인치까지 변경하였으며, 각 화면의 결과는 아래 <그림 4>와 같다. 실제 비디오 화면은 같은 콘텐츠를 사용하기에 비디오 콘텐츠에 제공된 해상도에 따라 화면의 비율이 제공되어, XML에서 정의한 가로와 세로의 비율과 정확히 일치 하진 않지만, 줄어드는 크기의 양은 비슷하며, 아래 <그림 4>는 초기 동영상 장면만을 캡처하여 7개의 화면을 재구성 하였다.

실제 DisplaySize adaptation을 통한 효과는 단말의 화

면 크기에 맞춤형 콘텐츠를 제공하는 것이 가장 큰 목적이라고 볼 수 있다. 이에 본 실험은 이러한 맞춤형 콘텐츠를 제공함으로써 화면 크기에 맞추어 재생되는 결과를 얻었다.



그림 12 minDisplaySize에 따른 화면 구성

더불어 큰 화면과 작은 화면을 사용했을 때 메모리 사용량을 비교 해보았다.

결과는 아래 그림 13에 나타나는 바와 같이 전체적으로 화면의 크기가 커질수록 많은 양의 메모리를 사용하는 것을 볼 수 있다.

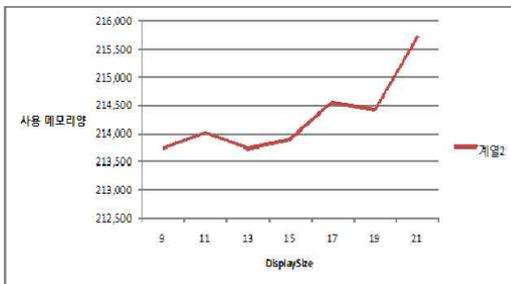


그림 13 DisplaySize에 따른 메모리 사용량

크기가 작은 화면을 재생하게 되면 사용되는 메모리가 작다는 것은 당연한 결과이지만, 같은 비디오 콘텐츠를 사용하기 때문에 이 실험의 결과는 비디오를 렌더링 하는데 사용되는 메모리량의 차이만을 보여 준다는 것을 알 수 있다.

즉 아직 실험을 수행하는 LAsER 재생기가 최적화 되지 않기 때문에 기본적으로 Player를 구동하는데 사용되는 메모리량이 많고, 이와 비교적으로 비디오의 렌더링에 필요한 메모리량의 크기 차이는 크지 않아 눈에 띄는 이득을 볼 수는 없지만 HD급의 TV에서와 보통 5인치 이하인 모바일 단말에서는 큰 차이를 보일 것으로 예상된다.

### 3. 결론

다양한 정보통신 단말기의 발전 및 왕성한 활용도를 통해 사용자의 멀티미디어 콘텐츠 소비 증대 및 이에 따라 다양한 단말 환경에 적합한 맞춤형 멀티미디어 서비스가 필요하다.

하지만 하나의 서비스로 단말기의 다양한 환경(전송 속도, 메모리 용량, CPU 속도, LCD크기 등)에 따라 맞춤형 서비스를 제공할 수 있는 서비스 규격은 아직 제정되지 않았고, 웹과 멀티미디어 관련 산업체와 표준 단체에서는 이에 대한 서비스 규격에 대해 활발히 활동하며 개발이 진행 중이다.

이 중 대표적인 멀티미디어 표준화 단체인 MPEG에서는 단말의 환경에 맞춤형 서비스를 개발하고자 하며, 사용자의 요구 사항을 반영한 양방향 멀티미디어 서비스에 적합한 규격인 LAsER를 통해서 이를 제공하고자 한다.

이에 본 논문에서는 이러한 기술의 기반인 장면 구성 기술에 대해서 W3C, MPEG의 표준에 대해서 알아보고, 그 중 MPEG의 LAsER 기술을 통해 적응적 장면 구성 방법인 Scene Adaptation을 접목시켰고, 실제 플레이어로 구현해 그 효율성에 대한 결과를 알아보았다.

본 논문에서 분석한 장면 구성 기술과 Scene Adaptation을 통한 연구 결과는 멀티미디어 콘텐츠 제작자가 의도하는 장면을 구성하도록 하여 콘텐츠 제작이 간편해지고, 콘텐츠 사용자는 멀티미디어를 시청할 수 있는 어떤 단말에서도 자신이 원하는 정보만을 제공받아, 향후 다양한 단말기를 대상으로 하는 리치미디어 콘텐츠가 더 원활하게 서비스 될 것이라는 효과를 기대한다.

현재 멀티미디어 관련 산업체와 단체에서는 웹에서 활발히 사용되는 widget 프로그래머, 멀티미디어에 widget과 같은 다양한 프로그램을 포함하여 서비스 하고자 하는 경향이 있다. 이에 본 논문에서 연구한 Scene Adaptation을 이러한 서비스에 적용하는 방법을 향후 연구 방향으로 남겨둔다.

#### 감사의 글

“본 연구는 지식경제부 및 정보통신산업진흥원의 대학 IT연구센터 지원사업의 연구결과로 수행되었음” (NIPA-2010-(C1090-1011-0001))

“지식경제부의 지식경제 기술혁신사업(“스케일러블 데이터 기반 Edutainment 핵심기술 개발”)의 연구결과로 수행되었음(과제번호 : 10035486)“

#### 참고 문헌

[1] 김규현 김병철 \*차지훈 “방송통신통합 환경에서 대화형 데이터 서비스를 위한 LAsER와 DIMS의 호환성 도출” 대한전자공학회 하계학술대회 2009. 07

[2] ISO, IEC 14496-20, Information technology, Coding of audio-visual objects. Part 20: Lightweight Application Scene Representation (LAsER) and Simple Aggregation Format (SAF)

[3] W3C SVG12, Scalable Vector Graphics (SVG) Tiny 1.2 Specification [Recommendation], <http://www.w3.org/TR/SVGMobile12>

[4] W3C SVG11, Scalable Vector Graphics (SVG) 1.1 Specification [Recommendation], <http://www.w3.org/TR/2003/REC-SVG11-20030114>