

논문-06-11-3-10

대화형 T-DMB 콘텐츠의 효율적인 전송을 위한 장면기술정보 최적화 기법

이 송 룩^{a)}, 차 경 애^{b)†}, 정 원 식^{c)}, 유 영 재^{c)}

An Optimization Technique of Scene Description for Effective Transmission of Interactive T-DMB Contents

Songlu Li^{a)}, Kyung-Ae Cha^{b)†}, Won-Sik Cheong^{c)}, and Yoo Young Jae^{c)}

요 약

디지털멀티미디어방송(Digital Multimedia Broadcasting, DMB)은 고품질의 방송 영상 및 부가 데이터를 이동 중인 차량 내에서도 모바일폰, PDA 등과 같은 휴대형 단말에서 수신할 수 있는 이동 멀티미디어 방송 서비스이다. 미디어 및 상호작용 정보의 인코딩 규격으로는 MPEG-4 표준을 사용하며, 대화형 콘텐츠의 사용자 상호작용 등의 정보는 장면기술정보인 MPEG-4 BIFS(Binary Format for Scene)에 의해서 이루어진다. 사용자와의 상호작용이 풍부하고 장면 구성이 복잡한 콘텐츠 일수록 BIFS 데이터가 전송되기 위해서 요구되는 비트율도 증가한다. 그러나 이동 단말에 고품질의 비디오 및 오디오 데이터를 전송하는 DMB 환경에서는 BIFS 스트림의 전송율은 매우 제한적이다. 한편 사용자 단말기에서 DMB 콘텐츠를 재생하려면 비디오, 오디오 등 미디어 객체들을 디코딩하기 전에 장면 구성 정보인 BIFS 스트림을 디코딩 및 파싱하여야 한다. 그러므로 MPEG-4의 높은 미디어 압축율에도 불구하고 제한된 대역폭과 장면 구성 정보 등 부가 데이터 전송 및 디코딩으로 인한 지연이 발생할 수 있다. 따라서 방송 콘텐츠의 특성 상 사용자의 임의 접근 시에도 재생될 콘텐츠의 BIFS 데이터를 파싱하여 대기 시간을 최소화하도록 초기 장면 구성 정보를 효율적으로 전송해야 한다. 본 연구에서는 지상파 DMB 대화형 콘텐츠의 장면기술정보인 BIFS 스트림을 저전송율 환경에 적응된 형태로 전달하기 위해서, 장면기술정보를 최적화하는 기법을 제안한다.

Abstract

The Digital Multimedia Broadcasting(DMB) system is developed to offer high quality audio-visual multimedia contents to the mobile environment. The system adopts MPEG-4 standard for the main video, audio and other media format. It also adopts the MPEG-4 scene description for interactive multimedia contents. The animated and interactive contents can be actualized by BIFS(Binary Format for Scene), the binary format for scene description that refers to the spatio-temporal specifications and behaviors of the individual objects. As more interactive contents are, the scene description is also needed more high bitrate. However, the bandwidth for allocating meta data such as scene description is restrictive in mobile environment. On one hand, the DMB terminal starts demultiplexing content and decodes individual media by its own decoder. After decoding each media, rendering module presents each media stream according to the scene description. Thus the BIFS stream corresponding to the scene description should be decoded and parsed in advance of presenting media data. With these reason, the transmission delay of BIFS stream causes the delay of whole audio-visual scene presentation although the audio or video streams are encoded in very low bitrate. This paper presents the effective optimization technique for adapting BIFS stream into expected MPEG-2 TS bitrate without any bandwidth waste and avoiding the transmission delay of the initial scene description for interactive DMB contents.

Key Words : T-DMB, BIFS, MPEG-4, Interactive multimedia, Scene description Optimization

a) 경북대학교 컴퓨터과학과
Department of Computer Science, Kyungpook National University
b) 대구대학교 정보통신공학부
School of Computer and Communication Engineering, Daegu University

c) 한국전자통신연구원 전파방송연구단 방송미디어연구그룹
Broadcasting Media Research Group, Radio and Broadcasting Research
Division, Electronics and Telecommunications Research Institute
† 교신저자 : 차경애(chaka@daegu.ac.kr)

I. 서론

지상파 디지털멀티미디어방송(Terrestrial Digital Multimedia Broadcasting, T-DMB)은 고품질 방송영상 등의 멀티미디어 콘텐츠를 차량 내에서도나 모바일폰, PDA 등과 같은 휴대형 단말에서 수신할 수 있는 방송통신 융합의 멀티미디어 서비스이다^[1-3].

T-DMB 콘텐츠는 장면을 구성하는 비디오, 오디오 또는 이미지 등의 다양한 형태의 미디어를 개별적으로 부호화하여 전송하며, 단말에서는 장면기술정보에 의해서 이들을 시공간적으로 구성하여 영상을 재생한다^[3,4]. 또한 장면기술정보는 대화형 콘텐츠의 사용자 상호작용을 위한 다양한 시나리오 정보를 기술한다. 즉 날씨, 주식과 같은 개별적인 부가 정보나 방송 영상 속의 상품 정보 등과 같이 영상 내용과 연동되는 정보를 함께 제공하여, 방송 시청 중에 메뉴 등을 클릭하여 해당 정보를 얻을 수 있게 함으로써 기존 방송 서비스와는 차별화되는 콘텐츠를 제공할 수 있다^[5-7]. 이러한 장면기술정보 역시 송신단에서 부호화하여 전송하는데, 이의 부호화된 데이터를 BIFS(Binary Stream For Scene)^[9-10]라고 한다. 이 때 실제 영상을 구성하는 미디어 스트림과 장면기술정보에서 기술하고 있는 객체의 연결 관계를 정의하는 데이터가 객체기술정보이며, OD (Object Descriptor)^[11]라고 한다. T-DMB 시스템에서 적용하고 있는 MPEG-4 장면기술정보는 Core2D 프로파일로 2차원 객체들의 컴포지션, 텍스처링, 로컬 인터랙션, 애니메이션, BIFS 업데이트, 웹링크 등의 다양한 2차원적 객체의 편집 및 구성 기능을 제공한다.

DMB 서비스가 상용화되면서, 점차 BIFS 데이터를 이용한 대화형 콘텐츠의 수요가 증가되고 있으며, 이미 다른 형태로 인코딩된 데이터 또는 기 제작된 DMB 콘텐츠를 재가공하여 사용할 경우에 같은 장면을 표현하더라도, 원 저작자에 따라서 서로 다른 형태의 장면기술정보를 가질 수 있다. 또한 사용자와의 상호작용이 풍부하고 장면 구성이 복잡한 콘텐츠일수록 장면기술정보의 양이 증가하면서 부호화된 BIFS 스트림의 전송율도 높아진다.

T-DMB 시스템에서 유효 데이터율은 64Kbps에서 1.152Mbps이며 그 중 비디오 스트림은 최대 544Kbps,

오디오는 최대 320Kbps, 이 외의 보조데이터는 최대 160Kbps의 유효 데이터율을 지원한다. 그러나 비디오 및 오디오 스트림에 대한 대역폭을 우선적으로 보장해야 하므로 실제로 OD/BIFS 스트림의 유효 데이터 전송율은 최대 32Kbps로 매우 제한적이다.

이와 같이 BIFS 정보의 전송은 다른 스트림에 비해 우선 순위가 낮고 대역폭도 제한된 반면에 사용자 단말에서 비디오, 오디오 등 미디어를 재생하기 위해서 BIFS 정보를 미리 디코딩하고 파싱하여 각 미디어 객체들의 시공간적 구성이 이루어져야 한다. 그러므로 MPEG-4 비디오 및 오디오 스트림의 높은 압축율에도 불구하고 BIFS 전송 및 디코딩으로 인한 지연이 발생할 수 있으며, 이로 인해 콘텐츠를 구성하고 있는 각 멀티미디어 객체의 디코딩 시간과 콘텐츠 재생도 지연될 수 있다. 따라서 사용자 단말에서의 초기 장면 지연을 단축하고 다양한 사용자 상호작용이 기술된 BIFS 정보를 제한된 수신 시간 및 가용한 대역폭에 효율적으로 전송할 수 있는 방안이 필요하다.

본 논문에서는 전송율 제한을 극복하고 효율적인 BIFS 데이터의 전송을 위해서 장면기술정보의 최적화 기법을 제안한다. 이를 위해서 MPEG-4 장면기술정보를 파싱하여 객체 및 장면의 구조를 파악한다. 이 결과를 바탕으로 장면기술정보의 일관성을 유지하면서, 삭제 또는 축소 가능한 요소를 선별하고, 장면을 재구성하여, 원래 정보와의 손실이 발생되지 않는 범위에서 최적화한다. 또한 최적화 과정에서 부호화된 BIFS 데이터의 바이트 수를 미리 예측하여, 가용한 비트레이트를 최대한 활용할 수 있도록 한다.

본 논문의 제 2절에서는 지상파 DMB 대화형 콘텐츠의 장면 구성 정보에 대하여 소개하고, 제 3절에서는 최적화를 위한 제안기법 및 시스템을 구성하는 각 모듈의 동작원리를 설명한다. 제 4절에서는 구현 및 실험 결과를 보이고, 제 5절에서 결론을 맺고 향후 연구방향을 제시한다.

II. 대화형 T-DMB 콘텐츠의 장면 구성 정보

다양한 미디어의 시공간적 시나리오 및 사용자 상호작용

정보를 기술하는 MPEG-4 장면기술정보는 VRML에 기반한 기술 형식을 가진다. 그림 1의 왼쪽은 장면기술정보를 기술한 텍스트의 일부이며, 오른쪽은 이의 재생 장면의 한 예를 보이고 있다.

그림 1에서 보이는 바와 같이 장면기술정보는 서로 다른 특성을 갖는 다양한 형태의 객체를 위치((1) Transform2D), 모양((2) Shape), 텍스처((3) Texture) 등의 정보를 기술하여 표현한다. 기하객체의 경우는 면 혹은 선의 스타일 등을 지정하는 정보((4) material), 객체의 종류를 나타내는 정보((5) geometry) 등으로 노드 등으로 구성되며, 이들은 속성을 나타내는 필드를 포함한다. 예를 들어 material의 경우는 면색을 나타내는 emissiveColor, 투명도를 결정하는 transparency 등의 속성 정보가 있다. 또한 각 노드는 고유 아이디를 지정할 수 있다. 그림 1의 (1)번의 경우

Transform2D 노드는 'I1_T2D'라는 아이디로 정의되었다.

사용자 상호작용 또는 애니메이션 등의 동적인 시나리오를 위해서 라우트(route) 노드를 이용한다. 라우트 정보에서는 마우스 클릭 등과 같은 이벤트의 주체가 되는 객체(소스 객체)와 그에 반응하는 대상 객체를 기술한다. 컨디션(Condition) 노드를 통해서 이벤트에 반응하는 구체적인 동작인 액션커맨드를 기술한다. 그림 1에서 사각형 객체를 클릭하였을 때 비로소 배경이 되는 이미지가 나타나는 시나리오를 기술하고자 한다면, 마우스 클릭 이벤트의 주체가 되는 사각형 객체에 터치센서(TouchSensor) 노드를 추가한다. 이 노드의 아이디를 'R1_TOS'로 정의한다고 가정하자. 또한 이미지 객체에 아이디가 'I1_CON_SHOW'인 컨디션 노드를 기술하여 클릭 이벤트가 발생하면, 해당 이미지가 나타나도록 기술한다. 이러한 상호작용을 위한 정보

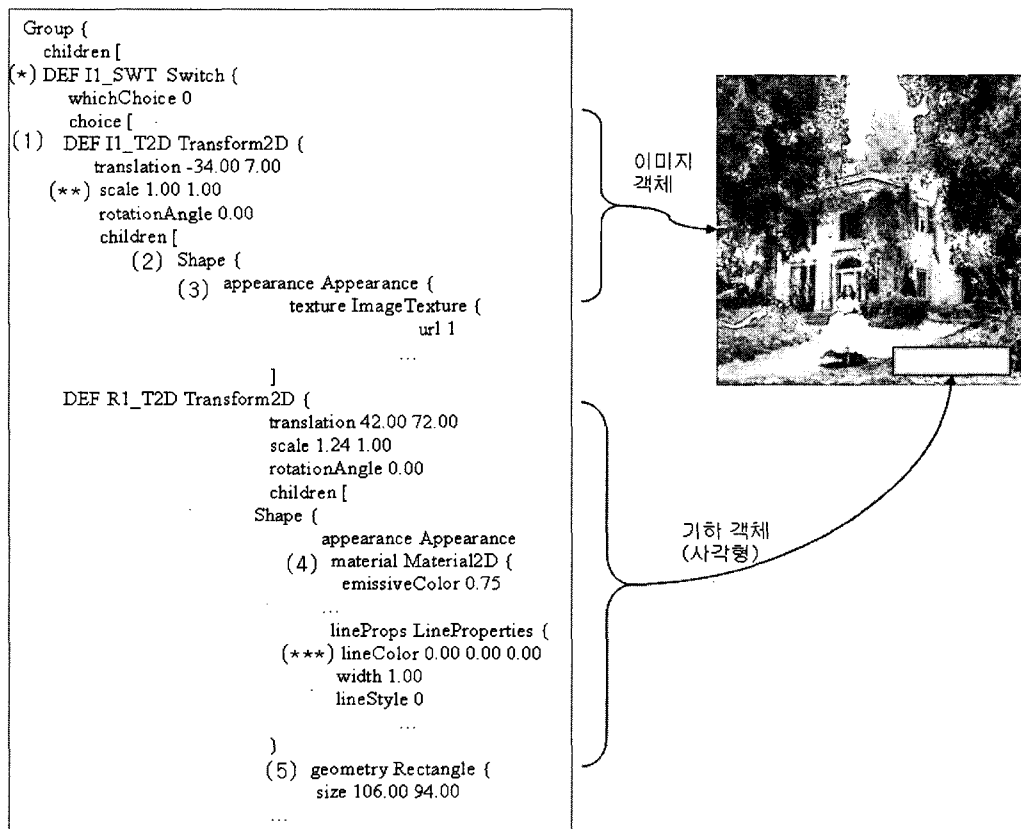


그림 1. 장면구성 정보와 객체의 재생 예
Fig. 1 A portion of Scene Description and its corresponding scene presentation

```

DEF R1_TOS TouchSensor { enabled true }
DEF I1_CON_SHOW Conditional {
    Buffer {
        (*) REPLACE I1_SWT.whichChoice BY 1 // 액션 커맨드
    }
}
ROUTE I1_TOS.isActive to I1_CON_SHOW.activate
    
```

그림 2. 대화형 시나리오를 위한 장면기술정보의 예
 Fig. 2 A portion of Scene Description for interactive scenario

는 그림 2와 같이 기술될 수 있다. 그림 1의 (*)에 정의된 스위치(Switch)노드는 객체가 화면에 보이거나 사라지게 하는 역할을 담당하며 그림 2의 (*)에 정의된 액션 커맨드에 의해서, 스위치 노드의 필드 값을 변화시켜서 이미지를 나타나게 하는 것이다.

그림 3은 위에서 설명한 장면기술정보를 노드 및 필드 단위로 구성한 트리 구조이다. 이와 같이 다양한 객체들을 기술하는 노드들은 계층적 구조를 이루며, 라우트 등의 노드를 통해서 상호관계를 표현한다. 따라서 최적화를 위한 재구성 시에 일관성 유지를 위해서, 장면을 이루는 시각적 객체를 파악해야 할뿐만 아니라, 이벤트 등의 상호 관계를 고려해야 한다.

Ⅲ. 대화형 T-DMB 콘텐츠를 위한 장면기술정보의 최적화

대화형 DMB 콘텐츠의 장면기술정보는 제작자에 따라서 같은 콘텐츠를 구성하더라도 다양하게 구성될 수 있다. 따라서 원래의 시공간적 정보를 손실하지 않고, 부호화 후 스트림의 바이트 크기가 감소된 형태로 최적화 할 수 있다. 이 절에서는 장면기술정보의 최적화 과정에 대하여 설명한다.

1. 장면 구성 정보의 분석

장면기술정보에서 객체를 기술하는 부분은 시각적 속성

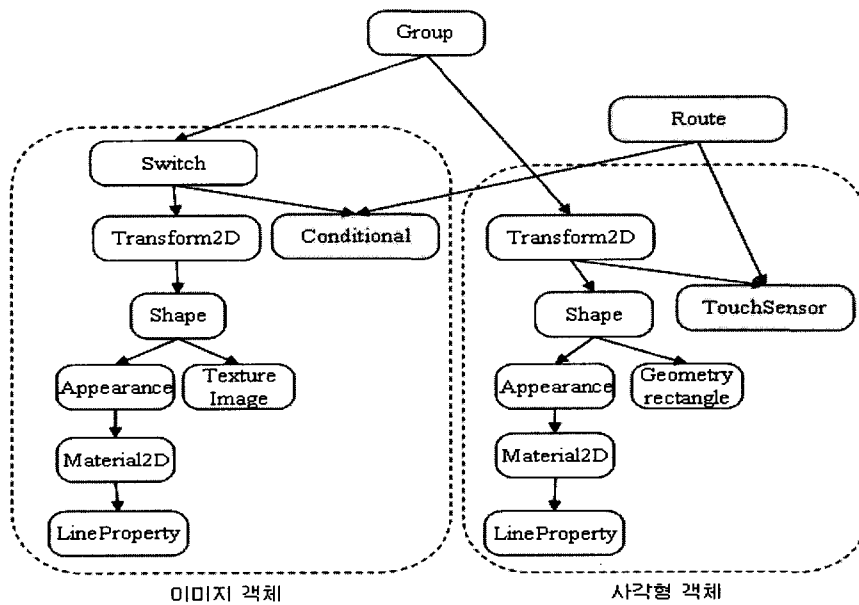


그림 3. 장면 구성 트리의 예
 Fig. 3 Example of Scene tree

을 표현하는 노드와 필드, 재생 시간을 나타내는 시간 정보 및 상호작용 정보를 표현하는 라우트 및 센서 노드들로 구성된다. 속성을 나타내는 노드들 중에는 그 필드의 값을 기술하지 않더라도 일반적으로 재생기에서 기본 값을 사용하여 재생 할 수 있는 특징을 가지는 요소들이 있다. 예를 들면 그림 1의 (**)-부분의 Transform2D 노드에서 scale 필드의 기본 값은 (1.00, 1.00)이며, (***) 부분의 LineProperty 에서 lineColor 필드의 기본 값은 (0.00, 0.00, 0.00)이다. 따라서 기본 값으로 기술된 필드를 선별하여 생략할 수 있다.

이와 같이 장면 구성 정보의 분석이 진행되면, 장면이 최적화되었을 경우 삭제되는 정보의 부호화 바이트 크기를 미리 파악 할 수 있으며, 이를 통해서 장면 구성 정보의 수정이 일어날 때마다 매번 부호화 바이트 크기를 검사해야 하는 오버헤드를 줄일 수 있다. 이를 위해서 각각의 정보 단위별로 텍스트가 부호화되었을 때 바이트수를 측정하여 장면기술정보 중 일부를 생략하였을 때 감소되는 BIFS 데이터의 바이트 수를 정량적으로 예측하도록 하였다.

이러한 BIFS 스트림의 바이트 수 예측을 위한 정보 수집 과정은 그림 4와 같다. 객체 노드 및 속성 필드의 단위별로 인코딩되는 바이트수를 측정하기 위해서 먼저 기존의 장면 기술정보를 파싱하여 멀티미디어 객체 별 속성 노드와 컨디션 노드, 라우트 노드 및 타임 노드를 분리하고, 그들의 상호 관계를 분석한다. 단위별 장면기술정보를 삭제할 때마다 부호화하고 이를 적용한 재멀티플렉싱을 통해서 기존 장면과의 변화를 확인한다. 이 때 일관성이 확인되면 전후의 BIFS 바이트 크기를 검사하고 이를 기록한다. 이러한 과정은 객체 노드, 속성 필드 및 속성값, 그룹 정보, 라우트, 시간, 컨디션 노드들에 대해서 실험하고 각 정보를 기록하여 최적화기의 바이트 크기 예측 모듈에서 활용하여 매번 장면기술정보를 인코딩하여 바이트 수를 검사하는 과정이 없이 가용한 자원에 최적화된 장면기술정보를 만들 수 있다.

이러한 과정은 BIFS 부호화 틀에 의해서 바이트 크기의 차이가 있을 수 있으나, DMB 장면기술 정보의 부호화는 MPEG-4 국제 표준화 그룹의 레퍼런스 소프트웨어 틀을 사용하였다. 그림 5는 기하객체 중 사각형을 기술하는 정보

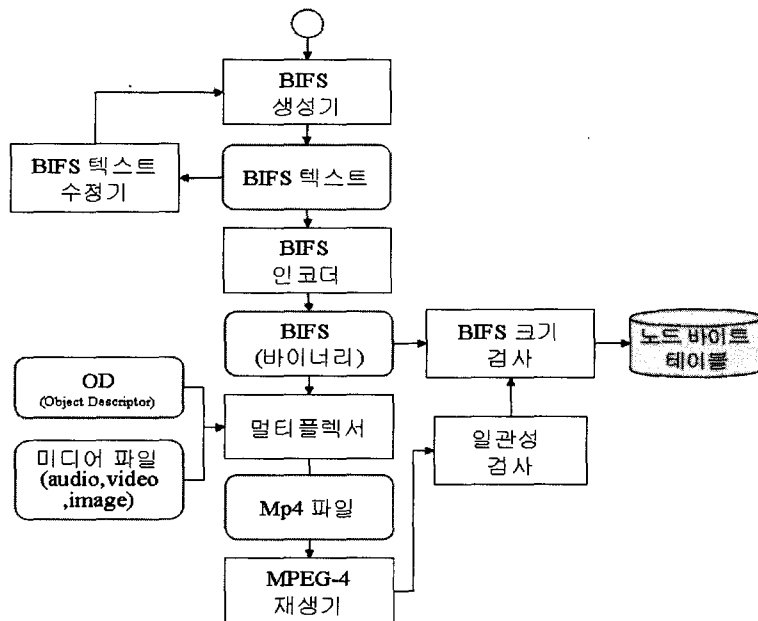


그림 4. BIFS 스트림의 바이트 수 예측을 위한 정보 수집 과정
 Fig. 4 Data collection process for BIFS stream

```

1) Group {
  2) children [
    3) DEF Switch3000 Switch {
      4) whichChoice 1
      5) choice [
        6) DEF Transform2D3000 Transform2D {
          7) translation 100.00 40.00
          8) scale 1.00 1.00
          9) rotationAngle 0.00
          10) children [
            11) Shape {
              12) appearance Appearance {
                13) material DEF Material2D3000 Material2D {
                  14) emissiveColor 0.00 0.10 0.50
                  15) filled TRUE
                  16) transparency -1.00
                  17) lineProps DEF LineProperties3000 LineProperties {
                    18) lineColor 0.00 0.40 0.30
                    19) width 1.00
                    20) lineStyle 0
                  }
                }
              }
            21) geometry Rectangle {
              22) size 106.00 94.00
            }
          ]
        }
      ]
    }
  ]
}

```

그림 5. 기하객체(사각형) 노드의 장면기술정보의 예
 Fig. 5 Example of geometry object(rectangle) description

를 보다 자세히 나타내고 이들의 부호화 바이트 수를 측정
 한 예이다.

만약, 부호화된 바이트 크기 정보가 파악되지 않은 노드
 나 필드를 기술하고 있는 장면기술정보가 입력되면, 파싱
 후 새로운 정보를 테이블에 추가해 나간다. 따라서, 실험과
 정을 거치면서, MPEG-4 BIFS에서 정의하는 거의 모든 노
 드 및 필드의 부호화 정보를 알 수 있다.

그림 5에서 사각형 객체의 정보를 최소로 표현하면,
 그림 6과 같이 나타낼 수 있다. 이 두 가지 경우의 장면
 기술정보를 재생했을 때, 사각형 객체의 변화는 그림 7
 과 같다.

그림 7의 (a)는 그림 5에 표현된 사각형 객체로 동영상의
 핫스팟을 나타내며, (b)는 그림 6의 장면기술정보에 의해서
 표현되는 사각형 객체를 보이고 있다. (b)에서 사각형 객체

```

DEF Transform2D3000 Transform2D {
  translation 100.00 40.00
  children [
    Shape {
      geometry Rectangle {
        size 60.00 40.00
      }
    }
  ]
}

```

그림 6. 기하객체(사각형) 노드의 장면기술정보의 최적화 예
 Fig. 6 Example of Scene Description optimization for geometry(rectangle) object

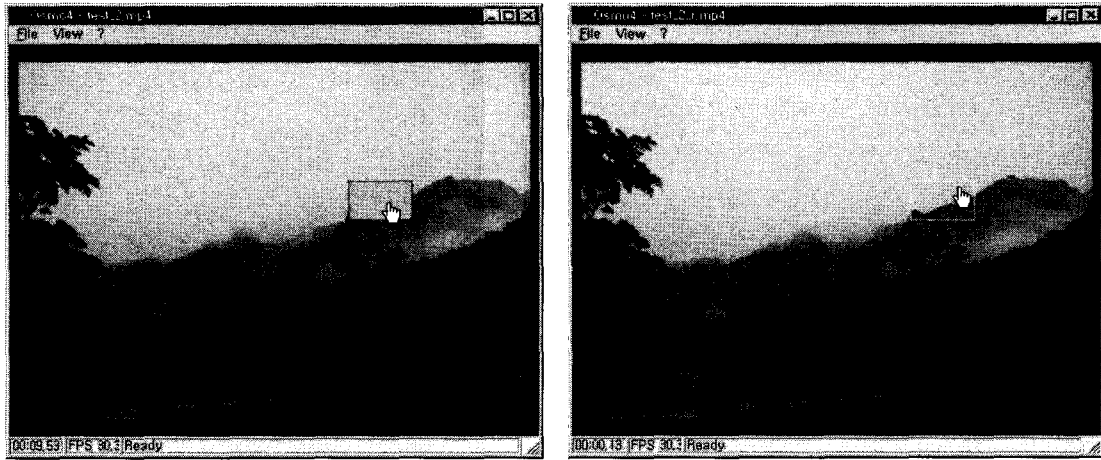


그림 7. 장면 재생의 예 : (a) 그림 5의 시각형 객체 정보로 핫스팟을 표현한 예 (b) 그림 6의 시각형 객체 정보로 핫스팟을 표현한 예
 Fig. 7 Example of Scene presentation : (a) Hot-spot presentation with scene description in Fig. 5 (b) Hot-spot presentation with scene description in Fig. 6

는 원래의 장면기술정보와 시각적 차이는 있으나, 핫스팟의 기능을 표현할 수 있으므로, 필요한 경우에 이와 같이 장면기술정보를 축소시키는 것이 가능하다. 본 논문에서는 이와 같은 기본 아이디어를 바탕으로 원래 장면 정보의 시각적 차이가 없는 범위에서 장면기술정보를 최적화하는 방안을 제시한다.

2. 장면기술정보 파싱

파서는 장면기술정보의 객체 및 시공간 시나리오를 파악하기 위해서, 그림 8과 같이 BIFS 텍스트 파일을 입력으로 하여 backspace와 '{', '}', '[', ']'와 같은 특수문자들을 기준으로 토큰을 생성한다. 토큰은 Reserved Simple 토큰, Related 토큰, Combination 토큰 등 세 가지로 정의한다.

Reserved Simple Token : 노드 파싱 테이블에서 매칭되는 문자열이다. 노드를 시작하거나 객체의 계층적 구조를 정의하는 키워드로 가장 심플한 토큰이다. 예를 들면, 'children', 'choice' 등과 같은 토큰들이다.

Related Token : 속성을 나타내는 값을 가지는 토큰들로서 숫자나 true/false 등의 할당된 값들이 연속적으로 나타나는 문자열이다. 예를 들면, 'emissiveColor', 'lineStyle', 'size' 등 필드 값을 가지고 있는 필드들이 Related Token에

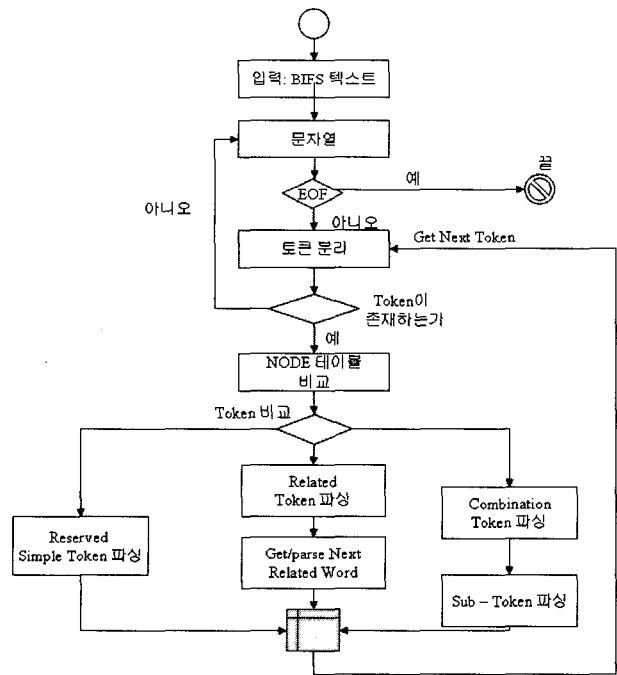


그림 8. 장면구성정보 파싱 모듈의 동작 과정
 Fig. 8 Processes for Scene Description parsing

속한다. 필드 값은 노드 파싱 테이블에 정의되지 않고 입력 값에 의해서 정해지므로 반드시 연관관계가 있는 필드와 같이 파싱하여야 한다.

Combination Token : 두 개 이상의 서브 토큰들의 조합으로 이루어진 토큰으로서 각 서브 토큰들은 '!'로 이어져 있다. **Combination Token**을 파싱 할 경우에는 각각의 서브 토큰을 **Related** 토큰으로 파싱하게 된다. **Combination Token**은 컨디션 노드나 라우트 노드, 시간정보 노드에서 찾아 볼 수 있다.

장면기술정보의 토큰이 분리된 후 객체 정보 및 사용자 상호작용, 시공간적 시나리오의 파악을 위해서 다음의 다섯 개의 컨테이너를 정의하고 각 정보를 구성한다.

- 객체 정보 컨테이너: 콘텐츠를 구성하고 있는 멀티미디어 객체의 속성 정보와 관련 센서 정보들을 저장한다. 객체 노드와 센서 노드의 연관성은 그들의 아이디를 통해서 구분한다. 멀티미디어 객체속성정보는 **Transform2D** 정보, **Material2D** 정보, **LineProperty** 정보 및 사용자 상호작용을 위한 **Touch Sensor** 등 각종 센서와 링크 정보를 포함한다.
- 라우트 정보 컨테이너: 라우트 노드에 기술된 이벤트의 주체가 되는 객체와 대상 객체의 노드 아이디, 액션 커맨드정보를 저장한다. 또한 이벤트가 적용되는 대상 객체에 기술된 컨디션 노드의 객체 아이디 및 액션

커맨드정보를 저장한다.

- 컨디션 정보 컨테이너: 멀티미디어 객체들 사이의 상호작용을 구현하기 위한 컨디션정보인 **Destination** 객체에 대한 액션 정보들을 저장한다. 주로 컨디션아이디, 액션 커맨드, **Destination**객체의 노드아이디 및 변경할 노드의 필드 값을 포함한다.
- 시간 정보 컨테이너: 멀티미디어객체의 시간정보를 저장한다. 주로 이벤트가 발생할 시간과 이벤트가 반영될 대상 객체의 노드아이디와 변경될 노드의 필드 값, 그리고 액션 커맨드를 포함한다.
- 업데이트 OD 정보 컨테이너: 콘텐츠에서 사용될 이미지, 비디오, 오디오 등 미디어 객체의 OD 정보를 저장하며 연관된 scr(script) 파일에서 객체의 미디어 속성을 구체적으로 기술한다.

위에서 정의한 컨테이너에 저장되는 정보들과 장면기술정보의 장면 구성 트리의 노드 및 필드와의 관계는 그림 9와 같다. 파싱된 결과를 바탕으로 각 노드와 필드를 분리하고, 그들의 아이디에 의해서 계층관계를 형성한다. 그러나, 장면 구성 트리의 모든 노드가 최적화 과정에 사용되는 정보가 아니므로, 필요한 정보들을 각각의 속성에 알맞은

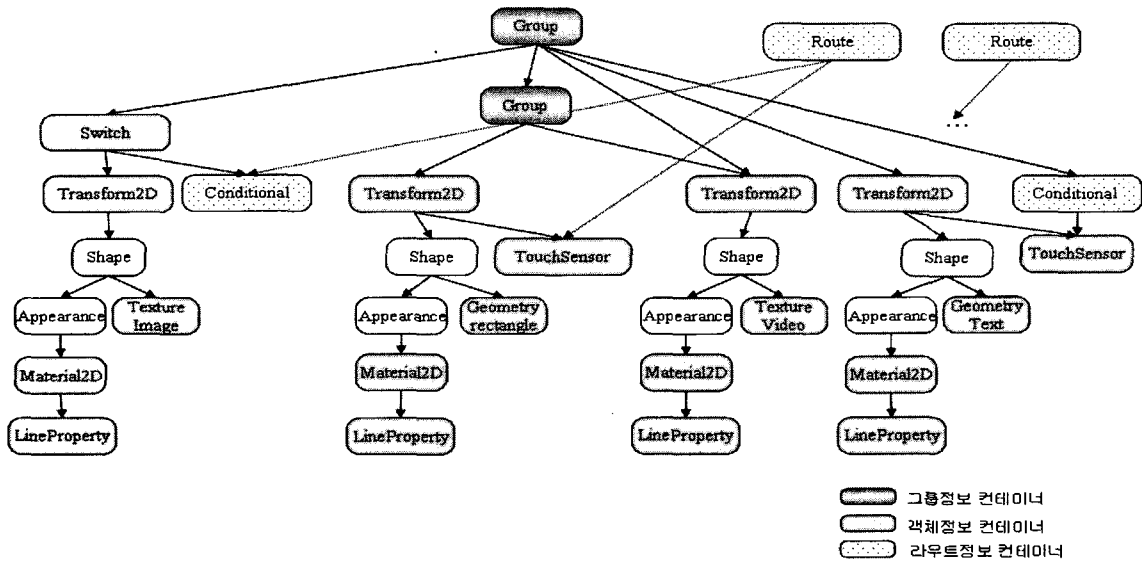


그림 9. 장면 트리의 노드와 컨테이너와의 관계
Fig. 9 Relation between Scene tree and containers

컨테이너에 저장하여 활용한다. 최적화 후 장면 구성 트리 에 그 결과를 반영하여, 트리를 수정하고 이를 기반으로 최 적화된 BIFS 텍스트를 생성한다.

3. 장면기술정보 최적화

최적화 모듈은 그림 10과 같이 최적화 프로세서, 장면 정

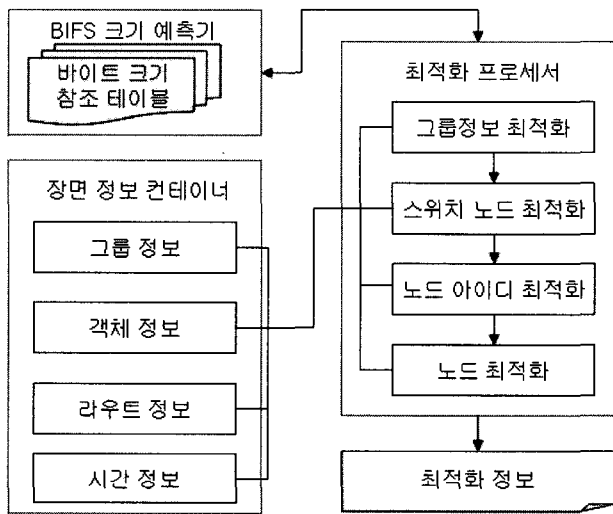


그림 10. 최적화 모듈 구성도
Fig. 10 Structure of Optimization modules

보생성기, BIFS 크기 예측기로 구성되며 BIFS 스트림이 목표 바이트 크기에 도달할 때까지 BIFS 바이트 크기 예측 모듈의 값을 참조하여 최적화 프로세서를 통하여 작업을 수행한다.

최적화 프로세서는 BIFS 텍스트 파서로부터 파싱한 장 면 구성 정보를 아래와 같은 세 가지 단계를 거쳐서 최적화 작업을 수행한다.

3.1 그룹 최적화

그룹 노드의 필요 여부를 판단하여 시공간적 장면 구성에 영향을 미치지 않는 노드에 대한 최적화 작업을 수행한다.

장면을 구성하는 여러 개의 객체들을 하나의 레이어 단 위로 묶어서 다룰 수 있는 기능을 제공하는 것이 그룹 노드 이다. 그룹 노드 내에는 또 다른 그룹 노드들을 정의할 수 있는데, 이 때 가장 큰 범위의 객체들을 포함하는 그룹 노 드를 루트 그룹이라고 하며, 그 내부의 그룹 노드들을 서브 그룹이라고 한다. 그림 11의 (a)는 텍스트 객체 하나와 이 미지 객체 하나를 포함하는 두 개의 서브그룹(Group1, Group2)이 정의되고, 이 두 서브그룹을 포함하는 상위 의 루트 그룹 노드(Group0)로 이루어진 장면을 예로 보이고 있다. 그리고 그림 11의 (b)는 이들의 계층 관계를 표현하 는 트리이다.

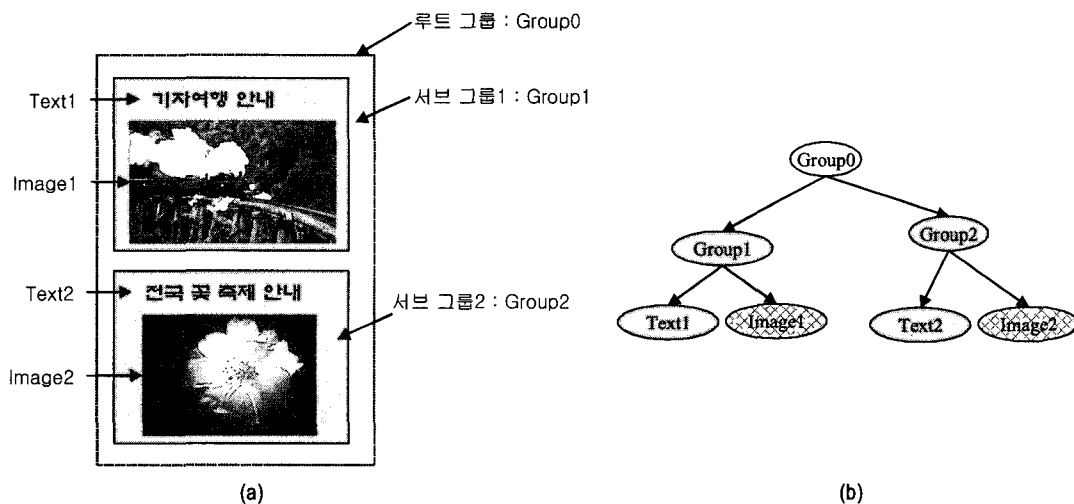


그림 11. 루트 그룹과 서브 그룹의 예 : (a) 장면 구성의 예 (b) 장면구성트리의 간략화된 표현의 예
Fig. 11 Example of Root Group node and Sub Group node : (a) Scene presentation example (b) Simple Scene tree example

서브 그룹은 객체의 시각적 재생에 영향을 미치지 않는 요소로서 삭제 가능하며 레이어 순서에 따라서 그룹에 포함된 객체만을 구성하여, BIFS 바이트 크기를 감소시킨다. 루트 그룹의 경우 정의된 아이디가 존재한다면 이를 삭제한다. 이 과정으로 아이디에 할당된 바이트 수만큼 BIFS 크기를 감소시키게 된다. 서브 그룹의 경우는 그 아이디를 참조하는 노드가 없다면 그룹 내의 객체들을 상위 그룹에 포함시킬 수 있으므로 해당 서브 그룹 노드를 삭제할 수 있다. BIFS 부호화 방식에 의해서 4 개의 서브 그룹을 단위로 그룹 노드를 삭제한 바이트 수에서 추가로 1 바이트씩 더 감소된다. 그림 12에서 (1)번의 최상위 그룹 아이디를 삭제하고, 서브 그룹 최적화에 의하여 서브그룹 'L0' 를 기술하고 있는 (2)번과 (3)번을 제외하였다. 이 과정으로 원래 BIFS 바이트 크기에서 총 13 바이트가 감소하게 된다.

3.2 스위치 노드 필터링

그림 1과 2에서 설명한 바와 같이 스위치 노드는 'choice' 필드의 값만으로 자신의 자식 노드가 화면에 보이거나 사라지게 하는 역할을 담당한다. 따라서 다른 노드와 달리 자식 노드들이 존재하더라도 라우트 노드, 컨디션노드 및 시간 정보 노드 등에서 그 아이디가 다시 사용되지 않는다면

삭제가 가능하다. 이를 위해서 객체별 최적화 작업을 수행하기 전에 스위치 노드의 존재 필요성을 검사한다. 라우트 정보 컨테이너에서 컨디션노드의 액션커맨드에 사용된 노드 아이디와 시간 정보 컨테이너에서 사용된 노드 아이디를 모두 비교하여, 같은 경우가 존재하지 않으면, 스위치 노드의 정의 부분을 삭제한다.

3.3 노드 아이디 최적화

노드 아이디는 저작자에 따라서 다양하게 표현할 수 있다. 그러나 노드 아이디는 문자 하나 당 1 바이트씩을 차지하므로, 최소의 문자수로 노드 아이디를 표현하면, 최적화에 매우 효율적이다. 노드 아이디는 그 첫 문자가 반드시 알파벳으로 시작하여야 하므로 첫 번째 문자에 사용할 수 있는 영문 대소문자('A/a'~ 'Z/z')를 글로벌 아이디라고 정의하여 할당하고, 한자리 정수('0'~ '9')를 로컬 아이디라고 정의하여 이들의 조합으로 이루어지는 두 문자의 아이디를 새로 할당한다. 글로벌 아이디는 객체를 구분하여 할당하고, 그 객체의 속성 노드는 로컬 아이디를 순서대로 할당한다. 따라서 최대 52개까지의 객체 아이디를 사용할 수 있으며, 객체의 속성 노드는 최대 10개까지 표현할 수 있다. 물론 이와 같은 방법은 아이디를 할당할 수 있는 객체 및 속성 노드 수의 한계가 있기 때문에 노드 수에 따라서 아이디 할

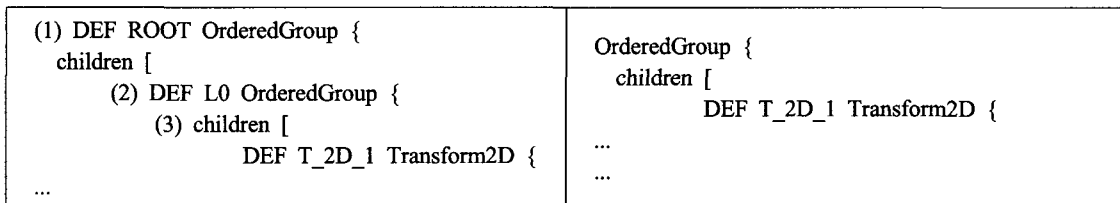
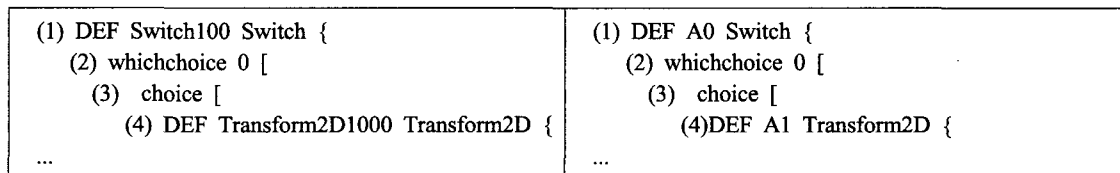


그림 12. 그룹 노드 최적화 예
Fig. 12 Example of Group node optimization



(a) (b)

그림 13. 노드 아이디 최적화 예
Fig. 13 Example of Node ID optimization

당을 위한 바이트 수를 증가시켜야 할 필요성이 있으나 객체의 식별 및 아이디 할당과정이 쉽고 일관적이다.

그림 13은 노드 아이디 최적화 과정을 수행한 후의 노드 아이디의 변화를 보여주고 있다. 객체의 아이디가 'A'로 할당된 경우, Switch 노드는 'A0', Transform2D 노드는 'A1', Material2D 노드는 'A2', 등 순서로 아이디가 부여되며 줄어든 문자 개수만큼 BIFS 바이트 수도 감소하였음을 알 수 있다. 그림 13의 (a)의 (1)번에서 (3)번까지는 스위치노드 정의 부분으로 18 바이트를 차지하며, (4)번의 경우도 역시 18바이트를 차지한다. (b)의 경우 (1)번에서 (3)번까지는 10바이트로, (4)번의 경우는 5바이트로 감소하였다.

컨디션 노드의 아이디는 객체 노드 아이디 및 객체 속성 노드 아이디와는 별도의 순서로 할당하기 위해서, 'z'에서 시작하여 알파벳의 역순으로 마지막 객체 아이디의 첫 문자와 일치하기 전까지 사용한다. 사용 가능한 문자가 없으면 'z' 문자에서부터 로컬 아이디를 두 자리로 늘린다. 콘텐츠를 구성하는 객체의 수가 50개이고 컨디션 정보의 개수가 40개라면 컨디션 노드에 대한 새로운 아이디는 표 1과 같이 부여된다.

표 1. 객체와 컨디션 노드의 아이디 할당 예
Table 1. Example of Node ID assignments

| 노드 타입 | 갯수 | 아이디 |
|-------|----|---|
| 객체 | 50 | 'A' ~ 'Z', 'a' ~ 'z' |
| 컨디션 | 40 | 'z0' ~ 'z9', 'y0' ~ 'y9', 'z00' ~ 'y19' |

객체 속성 노드와 컨디션 노드들에 대하여 아이디를 새로 부여한 후 그와 관련된 라우트정보와 시간정보에서도 해당되는 노드 아이디를 함께 변화시킨다. 이와 같은 아이디 할당 기법의 장점은 아이디만으로 객체의 구분이 가능하고 속성 노드들의 아이디 할당이 편리하며 노드의 첫 문자만 보더라도 어느 객체의 노드인지 분별이 가능하기에 검색 시간을 단축시킬 수 있다. 뿐만 아니라 각 노드들이 모두 같은 사이즈로 아이디를 부여하였기에 최적화 작업을 수행함에 있어서 감소될 BIFS 바이트 크기의 예측이 보다 편리하다.

3.4 노드 최적화

위에서 기술한 바와 같이 객체 속성 노드들 중에서 그 필드의 값을 기술하지 않더라도 재생기에서 기본 값을 사용하여 표현 할 수 있는 특징을 가지는 것들이 있다. 노드 최적화 과정에서는 이러한 기본 값들을 찾아내어 삭제함으로써 BIFS 바이트 수를 감소시킨다. 이를 위하여, 필드마다 해당 값을 삭제하였을 때 재생화면에 영향을 미치는지 여부를 검사하고 필드 값 삭제로 인해서 감소된 바이트 수를 검사함으로써 BIFS 사이즈 예측모듈을 위한 데이터를 구축하였다.

노드 최적화과정을 수행할 때는 반드시 레벨이 낮은 노드의 필드부터 삭제한다. 각 노드는 다른 노드를 자식 노드로 둘 수 있는데 레벨은 부모 노드일수록 높다. 자식 노드와 필드들을 삭제하기 전에는 부모 노드를 삭제할 수 없기 때문에 노드 최적화 과정은 반드시 필드부터 삭제하는 Bottom-Up 방식으로 진행 되어야 한다.

4. 장면 정보 재구성

최적화 프로세서를 통하여 최적화 과정을 거쳐 BIFS 바이트 크기의 예측치가 목표 값에 도달하였다면 수정된 장면기술정보를 BIFS 텍스트로 생성한다. 이 때 최적화 목표 값은 BIFS 전송을 위한 가용 자원의 크기 또는 사용자가 원하는 BIFS 부호화 스트림의 크기이다. 여기에서는 사용자 인터페이스 상에서 원래의 BIFS 바이트 크기보다 작은 값으로 사용자로부터 입력받은 값을 목표치로 한다.

BIFS 텍스트 재구성 모듈은 최적화 과정을 거쳐서 업데이트 된 각 컨테이너들에 저장된 수정 정보를 참조하여 삭제된 노드나 필드를 장면 정보 트리의 해당 트리 노드를 삭제한다. 그리고, 수정된 장면 정보 트리를 입력으로 하여 BIFS 텍스트를 생성한다.

IV. 구현 및 실험 결과

본 실험 환경은 Windows XP 운영체제이며, 인텔 펜티

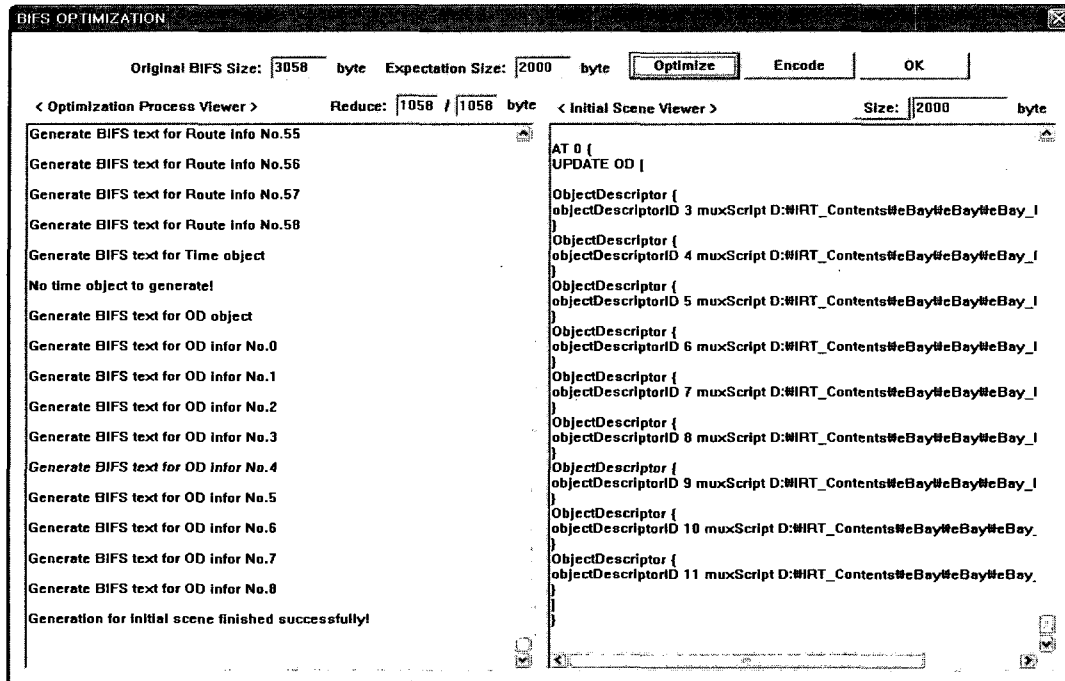


그림 14. 사용자 인터페이스
Fig. 14 User Interface

업 2GHz CPU와 2GB 램 (노트북)으로 구성된다. 시스템의 구현은 Microsoft Visual C++ 6.0을 이용하였으며 그림 14와 같은 사용자 인터페이스를 제공한다. 사용자 인터페이스의 최적화 프로세서 뷰어를 통해서 입력된 장면기술정보 텍스트의 단계별 최적화와 그에 따른 BIFS 크기의 변화 과정을 보여주고, 감소된 파일 사이즈와 감소해야 할 사이

즈를 예측하고 미리 계산한다. 재구성을 거친 장면 구성 정보는 초기 장면 뷰어를 통하여 확인할 수 있다. 생성된 초기 장면 정보는 인코딩 과정을 거쳐서 예측한 파일 사이즈와 비교할 수 있으며 멀티플렉싱 과정을 거쳐서 재구성된 콘텐츠를 검증한다.

표 2는 제안하는 최적화 기법의 테스트를 위하여 사용한

표 2. 샘플 콘텐츠의 장면기술정보
Table 2. Scene description for sample contents

| | 콘텐츠 1(eBay) | 콘텐츠 2 (JackieCola) | 콘텐츠 3(SmartPhone) |
|----------------|------------------|--------------------|-------------------|
| 원본 BIFS 데이터 크기 | 3058 bytes | 4069 bytes | 2724 bytes |
| 서브그룹의 수 | 8 | 10 | 8 |
| 객체 구성 | Video: 1 | Video: 1 | Video: 1 |
| | Audio: 1 | Audio: 1 | Audio: 1 |
| | Image: 7 | Image: 9 | Image: 7 |
| | Rectangle: 12 | Rectangle: 16 | Rectangle: 10 |
| | Total: 21 | Total: 27 | Total: 19 |
| 센서 노드 | Touch Sensor: 13 | Touch Sensor: 18 | Touch Sensor: 11 |

표 3. 각 단계별 최적화 후의 BIFS 바이트 크기의 변화

Table 3. Comparison of BIFS size : Original Scene Description & Optimized Scene Description

| BIFS 데이터 크기 \ 샘플 콘텐츠 | 콘텐츠 1(eBay) | 콘텐츠 2(JackieCola_Remake) | 콘텐츠3(SmartPhone) |
|----------------------|--------------------|--------------------------|--------------------|
| 원본 | 3058 bytes (100%) | 4069 bytes (100%) | 2724 bytes (100%) |
| 그룹 노드 최적화 수행 후 | 3001 bytes (98.1%) | 4000 bytes (98.3%) | 2667 bytes (97.9%) |
| 노드 아이디 최적화 수행 후 | 2196 bytes (71.8%) | 2925 bytes (71.9%) | 1961 bytes (72%) |
| 노드 최적화 수행 후 | 1664 bytes (54.4%) | 2251 bytes (55.3%) | 1450 bytes (53.2%) |

* Percentage (%) = (Optimized BIFS Size) / (Original BIFS Size) × 100

세 개의 샘플 파일에 대한 상세 정보이다. 각 콘텐츠의 장면기술정보는 모두 하나씩의 비디오와 오디오 객체를 기본으로 기술하고 있으며, 이미지와 기하객체, 및 터치 센서들로 대화형 콘텐츠를 위한 사용자 상호작용을 표현하고 있다. 뿐만 아니라, 서브 그룹 노드들을 포함하는 형태로 저작되었다.

표 3은 세 개의 샘플 콘텐츠들의 장면기술정보를 최적화하고 이를 부호화하여 BIFS 바이트 수의 변화를 측정 한 결과이다. 장면기술정보의 기술 방식에 따라서 최적화 과정을 거친 후 감소되는 바이트 크기는 다양하게 나타날 수 있다. 표 3에서와 같이 세 개의 샘플 콘텐츠는 그룹 최적화 후 원래의 BIFS 바이트 크기의 최저 98.3%, 최대 97.9%로 감소되었고, 노드 아이디 최적화 과정에서는 최저 72%, 최대 71.8%로 감소되었으며, 노드 최적화 과정에서는 최저 55.3%, 최대 53.2%로 감소되었음을 보여준다.

본 시스템에서는 최적화 과정에서 입력된 목표 값보다 BIFS 바이트 예측 값이 같거나 작아지면, 최적화 과정을 끝내어 원래 정보를 최대한 보장하고자 한다. 따라서 이 실험에서는 모든 최적화 과정을 수행하기 위해서 원본 BIFS 바이트 크기보다 매우 작은 값을 목표값으로 설정하였다. 실제로 입력된 각 BIFS 데이터의 바이트 크기의 50%에 해당되는 수치로 목표 값을 설정하여 실험한 결과이다.

이러한 실험 과정에서 BIFS 인코더의 프로세싱 시간을 측정하였다. 인코더의 BIFS 텍스트 파일을 open한 후

부호화하고 close 하기까지의 clock tick을 검사하고 이를 초 단위로 분석하였으며, 원본 파일과 최적화 결과 파일을 각각 5회씩 인코딩하여 프로세싱 시간의 평균 및 최대, 최소값을 계산하였다. 평균 프로세싱 시간은 0.093초이며, 최대 및 최소 프로세싱 시간은 각각 0.14초와 0.079초로 계산되었다. 이와 같은 결과는 컴퓨터 시스템의 환경에 따라서 달라질 수 있을 것이므로, 본 논문에서는 실험의 결과를 BIFS 데이터의 전송율에 적용하여 보고자 한다. 서론에서 언급한 바와 같이 BIFS 데이터의 전송율이 32Kbps이고 본 실험에서 바이트 크기가 가장 큰 파일인 콘텐츠 2(4kbyte)를 인코딩하는 평균 시간을 실험 결과에 의해서 0.093초라고 한다면, 32 kbyte를 인코딩하는데는 0.744초가 걸린다고 예상할 수 있다(0.093초 x 32kbyte / 4kbyte).

그림 15 (a)는 샘플 콘텐츠 1의 원본 장면기술정보의 일부를 보여주는 텍스트이다. 이 콘텐츠 기술정보에서 사운드 노드의 'AudioSource'의 여러 가지 필드 중 기본 재생에 영향을 미치지 않는 요소들과, 서브그룹 노드, 노드 아이디 재할당등을 통해서 최적화된 장면기술정보 텍스트의 일부를 그림 15의 (b)에서 보이고 있다.

그림 16은 샘플 콘텐츠 1의 최적화 전과 최적화 후의 재생 화면을 보여주고 있다. 제안하는 최적화 기법이 장면기술정보에 대한 무손실 최적화 기법이므로, 최적화 전과 최적화 후의 콘텐츠를 동일한 재생 시간에 동일한 이벤트를 적용시켰을 때의 화면을 캡처한 결과, 재생화면 역시 동일함을 보여준다.

```

16 )
17 texture MovieTexture (
18 loop false startTime 0.000000 stopTime 398.9333191 url 3
19 )
20 )
21 )
22 )
23 )
24 Sound2D {
25 source AudioSource {
26 pitch 1.0000000 speed 1.0000000 startTime 0.0000000 stopTime -1.000
27 }
28 }
29 )
30 )
31 DEF L1 OrderedGroup (
32 children [
33 DEF I3 SWI)Switch (whichChoice 0
34 choice [
35 DEF I3_T2D Transform2D (scale 1.0000000 1.0000000
36 translation -122.0000000 -92.0000000
37 children [
38 Shape {
39 geometry Bitmap {
40 scale 1.0000000 1.0000000
41 }
42 appearance Appearance {
43 material DEF I3_M2D Material2D (emissiveColor 0.0000000 0.0000000 (
44 filled true
45 transparency 0.0000000
46 )
47 texture ImageTexture (
                    
```

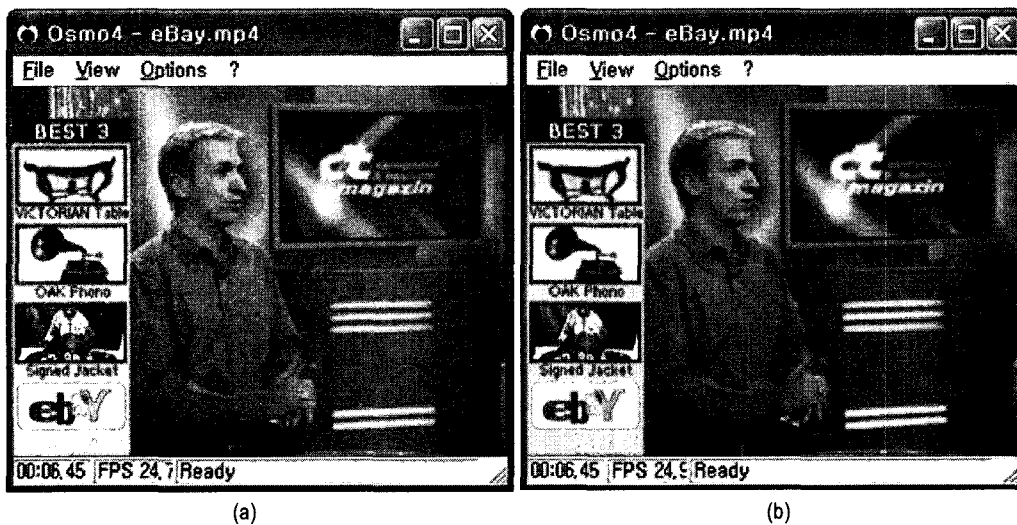
```

13 )
14 texture MovieTexture (
15 url 3
16 )
17 )
18 geometry Bitmap {
19 scale 1.00 1.00
20 )
21 )
22 ]
23 )
24 Sound2D {
25 source AudioSource {
26 url 4
27 }
28 )
29 DEF B0 Switch (
30 whichChoice 0
31 choice [
32 DEF B1 Transform2D (
33 translation -122.00 -92.00
34 children [
35 Shape {
36 appearance Appearance {
37 material DEF B2 Material2D (
38 emissiveColor 0.60 0.00 0.00
39 filled TRUE
40 transparency 0.99
41 )
42 texture ImageTexture (
43 url 5
44 )
                    
```

(a)
(b)

그림 15. 샘플 콘텐츠 1의 장면기술정보 : (a)원본 장면기술정보 (b)최적화 후 장면기술정보

Fig. 15. Example of Scene Description for Sample Content 1: (a) Original Scene Description (b) Optimized Scene Description



(a) (b)

그림 16. 샘플 콘텐츠1의 재생 장면 : (a)원본 BIFS를 이용한 장면 (b)장면기술정보 최적화 후 재생 장면

Fig. 16 Example of Scene presentation for Sample Content 1: (a)Scene presentation with Original BIFS (b) Scene presentation with optimized BIFS

VI. 결론 및 향후 연구방향

본 연구에서는 대화형 서비스를 위해 효율적으로 지상파 DMB 콘텐츠의 장면기술정보를 최적화하는 기법을 제시하였다. 실험 결과가 보여 주듯이 제안하는 최적화 작업을

모두 수행하였을 때 장면구성 정보는 최소 55.3%, 최대 53.2%로 축소되었다. 다양한 방식으로 기 저장된 콘텐츠를 DMB 대화형 방송용 콘텐츠로 변환하거나 제한된 대역폭에서 보다 효율적인 장면정보의 전송에 활용한다면 DMB 상에서의 데이터방송을 위한 응용 콘텐츠를 보다 효율적으

로 서비스할 수 있을 것이다. 또한 방송 콘텐츠의 특성 상 사용자의 임의 접근 시에 제한된 시간 내에 초기 장면을 재생하고 전송 지연을 방지하는 기반 기술로 활용될 수 있다.

앞으로의 연구 방향은 보다 효율적인 장면정보 최적화를 위해서, 초기 장면에서 객체의 재생 시점을 한 BIFS 커맨드를 이용한 장면정보전송 기법을 활용하는 것이다.

참 고 문 헌

[1] G. Lee, S. Cho, K. Yang, Y. Hahm and S. Lee, "Development of Terrestrial DMB Transmission System based on Eureka-147 DAB System." IEEE Transactions on Consumer Electronics, Volume 51, Issue 1, pp.63 - 68, Feb. 2005.

[2] V. Ha, S. Choi, J. Jeon, G. Lee, W. Jang and W. Shim, "Real-time Audio/Video Decoders for Digital Multimedia Broadcasting," Proceedings of the 4th International Workshop on System-on-Chip for Real-Time Applications, 2004.

[3] 초단파디지털라디오방송 데이터송수신정합표준, 한국정보통신기술협회, 2005.

[4] "DMB/IPTV/DCATV 기술워크샵" 발표자료집, 대한전자공학회, 3. 2005.

[5] 김상훈, 박찬섭, 김만식, "T-DMB 실시간 비디오 부가데이터 서비스 시스템 개발", 한국방송공학회논문지, 제10권 4호 통권29호, pp.474-487, 2005.

[6] 데이터 방송 서비스 국내외 현황, 한국전자통신연구원, 2004.

[7] 맞춤형 방송규격, 정보통신단체표준, 2004. 11.

[8] J. Signes, Y. Fisher, and A. Eleftheriadis, "MPEG-4's binary format for scene description," Signal Processing: Image Communication, vol.15, issues 4-5, pp. 321-345, Jan. 2000.

[9] Olivier Avaro, Alexandros Eleftheriadis, "MPEG-4 Systems: Overview" Signal Processing: Image Communication, vol.15, issues 4-5, pp. 281 - 298, Jan. 2000.

[10] WG11 (MPEG) MPEG-4 Overview (V.21 - Jeju Version) document, ISO/IEC JTC1/SC29/WG11 N4668, Mar. 2002.

[11] C. Herpel and A. Eleftheriadis, "MPEG-4 Systems: elementary stream management," Signal Processing: Image Communication, vol. 15 no. 4-5, pp. 299-320, Jan. 2000.

저 자 소 개



이 승 록

- 2001년 : 중국 길림성 연변과학기술대학 (학사)
- 2005년 : 경북대학교 컴퓨터학과 (석사)
- 2005년~현재 : 경북대학교 컴퓨터학과 (박사과정)
- 주관심분야 : Multimedia System, Visual Languages, and Embedded Software 등



차 경 애

- 1996년 2월 : 경북대학교 컴퓨터학과 (이학사)
- 1999년 2월 : 경북대학교 컴퓨터학과 (이학석사)
- 2003년 8월 : 경북대학교 컴퓨터학과 (이학박사)
- 2001년 3월 ~ 2004년 2월 : 경북대학교 초빙교수
- 2004년 3월 ~ 2005년 2월 : 한국정보통신대학교 연구교수
- 2005년 3월 ~ 현재 : 대구대학교 정보통신공학부 전임강사
- 주관심분야 : 멀티미디어시스템, 디지털방송, 콘텐츠 저작, 대화형 미디어 등

— 저 자 소 개 —



정 원 식

- 1992년 2월 : 경북대학교 전자공학과 (공학사)
- 1994년 2월 : 경북대학교 대학원 전자공학과 (공학석사)
- 2000년 2월 : 경북대학교 대학원 전자공학과 (공학박사)
- 2000년 5월 ~ 현재 : 한국전자통신연구원 전파방송연구단 방송미디어연구그룹 대화형미디어연구팀 선임연구원
- 주관심분야 : 영상처리 및 압축, 멀티미디어 시스템, 대화형방송, DMB, MPEG 등



유 영 재

- 1998년 : 충남대학교 컴퓨터공학과 (공학사)
- 2000년 : 충남대학교 컴퓨터공학과 (공학석사)
- 2004년-현재 : 한국전자통신연구원 대화형미디어연구팀 연구원
- 주관심분야 : 디지털 대화형 방송, 지상파DMB 시스템, MPEG-4 시스템 등