

전송율을 고려한 지상파 DMB 대화형 콘텐츠의 MPEG-4 씬 디스크립션 최적화 기법

차경애¹, 조지연²⁰, 이송록³, 김상욱⁴, 유영재⁵, 정원식⁶

¹⁾ 대구대학교 정보통신공학부, ^{2,3,4)} 경북대학교 컴퓨터과학과, ^{5,6)} 한국전자통신연구원

¹⁾chaka@daegu.ac.kr, ^{2,3)}{jycho, songlock}@woorisol.knu.ac.kr, ⁴⁾swkim@cs.knu.ac.kr

Optimization Technology for MPEG-4 Scene Description of Interactive T-DMB contents

Kyung-Ae Cha¹, Jiyeon Cho²⁰, Songlu Li³, Sangwook Kim⁴, YuYoung Jae⁵, Won-Sik Cheong⁶

¹⁾School of Computer and Communication Engineering, Daegu University

^{2,3,4)}Department of Computer Science, Kyungpook National University

^{5,6)}Electronics and Telecommunications Research Institute

요 약

이동멀티미디어 방송 서비스의 국내 표준인 디지털 멀티미디어 방송(DMB) 시스템에서 대화형 콘텐츠는 MPEG-4 씬 디스크립션 정보에 의해서 이루어진다. 따라서 사용자 상호작용이 풍부한 콘텐츠일수록 씬 디스크립션이 전송되기 위해서 요구되는 비트레이트도 높아진다. 그러나 이동 단말에 고품질의 멀티미디어 데이터를 전송하는 지상파 DMB 환경에서는 씬 디스크립션과 같은 부가 데이터의 전송율을 매우 제한적이다. 그러므로 오디오나 비디오 데이터뿐만 아니라 씬 디스크립션 정보도 저대역폭 및 이동 단말을 대상으로 유효한 품질을 보장할 수 있는 형태로 인코딩되어야 한다. 본 논문에서는 콘텐츠를 전송하기 전에 인코딩된 MPEG-4 씬 디스크립션 정보를 파악하여 전송시점에 예상되는 비트율에 최적화되도록 재구성함으로써 씬 디스크립션 정보의 전송 지연 및 손실을 최소화하는 방안의 연구 내용을 제안하고 실험결과를 소개한다.

1. 서 론

이동 환경에서 고품질의 방송 콘텐츠 및 교통 정보나 상품 정보 등의 부가 데이터를 전송하는 이동 멀티미디어 시스템인 지상파 DMB는 미디어 및 상호작용 정보의 인코딩 규격으로 MPEG-4 표준을 사용한다[1-3]. 따라서 사용자 상호작용 정보를 포함하는 대화형 콘텐츠를 서비스하기 위한 MPEG-4 씬 디스크립션 정보를 인코딩하여 미디어와 함께 전송한다. 이 때 상호작용이 풍부한 콘텐츠일수록 씬 디스크립션이 전송되기 위해서 요구되는 비트레이트도 높아진다. 그러나 이동 단말에 멀티미디어 데이터를 전송하는 지상파 DMB 환경에서는 씬 디스크립션과 같은 부가 데이터의 전송율을 매우 제한적이다. 또한 방송 콘텐츠의 특성상 사용자의 임의 접근 시에도 재생될 콘텐츠의 씬 정보를 파싱하여 대기 시간을 최소화하도록 초기 씬 디스크립션 정보를 전송해야 한다. 그러므로 오디오나 비디오 데이터뿐만 아니라 씬 디스크립션 정보도 저대역폭에 유효한 품질을 보장할 수 있도

록 인코딩될 필요가 있다.

본 논문에서는 콘텐츠를 전송하기 전에 인코딩된 MPEG-4 씬 디스크립션 정보를 파악하여 전송시점에 예상되는 가용한 비트율에 최적화되도록 재구성함으로써 씬 디스크립션 정보의 전송 지연 및 손실을 최소화하는 방안을 제안한다.

MPEG-4 씬 디스크립션에 기술되는 각 객체는 그 특성 별로 스크립트 및 코딩된 스트림인 BIFS(Binary Format for Scene)의 비트레이트가 상이하다. 따라서 인코딩된 원본 씬 디스크립션에서 객체 별 속성 및 씬 구조를 파악하고 가용 전송율을 초과하지 않으면서 원래의 씬의 정보 손실을 최소화하도록 초기 씬을 구성하는 객체의 종류 및 객체 속성 등을 재구성하여 인코딩한다. 따라서 초기 씬의 정보가 지연이나 손실 없이 전송될 수 있다. 이후 초기 씬에 포함되지 못한 정보는 해당 객체 및 상호작용 정보의 재생 시간에 맞추어서 BIFS 커맨드로 전송하여 제한적인 저전송환경에 씬 디스크립션 정보를 효율적

으로 전송한다.

본 논문은 제 2절에서 지상파 DMB 시스템에 대하여 설명하고, 제 3절에서는 MPEG-4 콘텐츠의 씬 디스크립션 구조 및 특성을 설명한다. 제 4절에서는 대화형 MPEG-4 콘텐츠의 최적화 기법에 대하여 구체적으로 설명하고 최적화된 실험 결과를 보이며 제 5절에서 결론을 맺는다.

2. 지상파 DMB에서의 콘텐츠 서비스 표준

지상파 DMB는 유럽의 디지털 오디오 방송(Digital Audio Broadcasting, DAB) 표준에서 규정한 스트림 모드를 통하여 MPEG-4 비디오(Advanced Video Coding, AVC) 데이터, MPEG-4 오디오 (Bit-Sliced Audio Coding, BSAC) 데이터 및 대화형 데이터 방송을 위한 MPEG-4 BIFS(Binary Format for Scenes) 데이터를 MPEG-4 SL(Sync Layer)와 MPEG-2 TS(Transport Stream)로 다중화한 후, 오류 보호 매커니즘이 적용된 스트림을 전송한다. 이 과정은 그림 1과 같이 나타낼 수 있다.

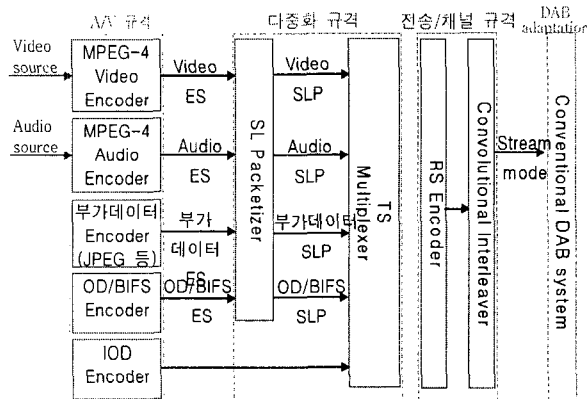


그림 1. 지상파 DMB 콘텐츠 인코딩 과정

전송 스트림 포맷인 MPEG-2 TS 패킷은 그림 2와 같이 패킷 하나가 188 바이트의 일정한 크기로 이루어지며 헤더를 제외한 184 바이트의 페이로드를 가진다[4]. MPEG-4 씬 디스크립션 정보는 인코더에 의해서 BIFS ES로 생성되며, 이 스트림은 액세스 유닛(Access Unit, AU)단위로 SL 패킷화 및 패킷화된 ES(Packetized ES, PES)로 MPEG-2 TS의 페이로드에 실리게 된다. 이 때 BIFS 데이터가 탑재될 수 있는 가용한 페이로드의 크기는 전송 시점의 비디오나 오디오 또는 이미지 데이터의 비트레이트에 의해서 변화된다. 그러므로 TS 패킷화과정에서의 가용한 페이로드 크기에 적합한 비트레이트의 BIFS 데이터를 생성하여, 가변 전송율에 적응적으로 콘텐츠를 전송할 수 있다.

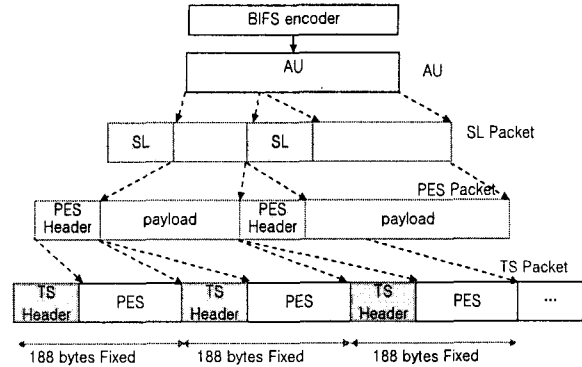


그림 2. BIFS 데이터의 MPEG-2 TS 패킷화 과정

3. MPEG-4 씬 디스크립션 분석

다양한 미디어의 시공간적 시나리오 및 사용자 상호작용 정보를 기술하는 MPEG-4 씬 디스크립션은 VRML에 기반한 텍스트 기술형식을 가진다. 그림 3은 기하객체인 사각형을 표현하는 씬 디스크립션의 예의 일부이다. 그리고 각 텍스트가 인코딩되어 부호화되었을 때 바이트수(byte)를 표시하였다. 이와 같이 각 기하 객체 및 미디어 객체를 나타내는 스크립트, 사용자 상호작용 정보 및 객체의 시간정보를 나타내는 스크립트 별로 부호화된 바이트수를 측정하였다. 이로써 스크립트 중 일부를 생략하여 BIFS의 크기를 축소할 때, 정량적으로 감소되는 바이트수를 예측할 수 있다. 예를 들어 그림 3에서 14번째 줄의 사각형 객체의 색상 정보를 생략한다면, 원래의 BIFS 데이터에서 4 바이트가 감소된다.

```

1) Group {
2) children [
3) DEF Switch3000 Switch {
4)   whichChoice 1
5)   choice [
6)     DEF Transform2D3000 Transform2D {
7)       translation -187.00 142.00 -----8바이트
8)       scale 1.00 1.00 -----8바이트
9)       rotationAngle 0.00 -----4바이트
10)      children [
11)        Shape {
12)          appearance Appearance { -----1바이트
13)            material DEF Material2D3000 Material2D {
14)              emissiveColor 0.75 0.75 0.75
15)              filled TRUE
16)              transparency -1.00
17)              LineProperties {
18)                lineColor 0.00 0.00 0.00
19)                width 1.00
20)                lineStyle 0
                }
14) geometry Rectangle { -----1바이트
22)   size 106.00 94.00 -----8바이트
            }
          }
        }
      ]
    }
  ]
}
    
```

그림 3. MPEG-4 씬 디스크립션의 예

4. MPEG-4 씬 디스크립션 최적화

본 절에서는 BIFS 데이터를 가변적인 저전송율에 효율적으로 전송되도록 재구성하는 최적화 방법을 설명한다.

그림 3에서 볼 수 있듯이 객체를 기술하는 스크립트는 선택, 내부색, 투명도 등의 정보를 표현하는 부분, 스케일 및 모양을 표현하는 부분 등으로 나누어진다. 또한 재생 시간을 나타내는 시간 정보 및 상호작용 정보를 표현하는 라우트 및 센서 노드들로 추가하여 다양한 씬을 구성할 수 있다. 본 논문에서는 씬 디스크립션 내의 개별적인 노드 및 노드 내의 필드를 구분하여, 초기 씬의 시각적 정보를 최대한 보장하는 형태로 BIFS정보를 축소한다. 이는 사용자가 최초로 서비스 받는 콘텐츠의 원형을 최대한 보장하기 위함이다.

이를 위해서 원래의 씬 디스크립션 내의 정보 중 초기 씬에서 제외될 부분에 대한 우선순위 선정이 필요하다. 여기에서는 객체 노드의 필드 중에서 표시하지 않아도 재생 가능한 기본 값들을 먼저 생각한다. 그 후, 가용한 비트율과의 비교를 계속하면서, 초기 씬에 서술된 상호작용 정보인 라우트(Route) 및 센서 노드들을 생각한다. 그러나 이 정보는 라우트의 소스 및 센서 노드와 연결된 객체 노드가 재생되는 시점에서 BIFS 커맨드로 재생성되어 전송된다. 상호작용 정보의 생략 후에도 가용한 비트율을 초과하는 경우는 객체 노드를 생각한다. 이 때, 씬 전체의 시간 시나리오를 파악하여 초기 씬에 반드시 포함되어야 하는 객체 및 상호작용 정보의 생략으로 재생되지 않는 객체를 선별하여 전자는 포함하고, 후자는 생략하는 방법을 사용한다.

이와 같은 방법으로 그림 3에서 제시되었던 BIFS 텍스트는 그림 4와 같이 최적화 될 수 있다.

```

1) Group {
  2) children [
    6) DEF Transform2D3000 Transform2D {
      10) children [
        11) Shape {
          21) geometry Rectangle {
            22) size 187.00 145.00
          }
        }
      ]
    }
  ]
}
    
```

그림 4. 그림 3의 최적화된 MPEG-4 씬 디스크립션 예

표 1에서는 최적화 작업을 수행한 각 단계별 BIFS 바이트 수를 보여준다. 원본 파일 사이즈는 총 199바이트의 사이즈를 가진다. 그 중 TimeSensor를 삭제하면 83바이트가 줄어들었고 Default 값들을 삭제하면 136바이트가 줄어들었다. Route값을 정보를 삭제했을 경우 161바이트가 줄어들었으며 Location정보 삭제 시에는 170바이트가 줄어들었다. 마지막으로 Attribute 정보를 삭제하면 총 173바이트가 줄어들어 전체의 13.07%만큼의

바이트수를 가진다. 기본값을 삭제하는 단계까지는 콘텐츠의 변화가 없으며 그 이상으로 최적화를 진행하였을 경우 BIFS 커맨드를 이용하여 콘텐츠를 업데이트해야 한다.

표 1 초기 씬의 최적화 과정의 바이트 수 비교

BIFS 바이트수	감소된 바이트수	%	Remark
199	----	100%	Full BIFS
116	83	58.29%	Delete Time Sensor
63	136	31.66%	Delete Default Values
38	161	19.10%	Delete Route Info
29	170	14.57%	Delete Location Info
26	173	13.07%	Delete Color Attribute

5. 결론 및 향후 연구과제

본 연구에서는 대화형 서비스를 위해 효율적으로 지상파 DMB 콘텐츠를 저작하고 최적화하는 기법을 제시하였다. 실험 결과가 보여 주듯이 제안하는 최적화 작업을 수행하였을 때 장면구성정보는 최소 58.59%, 최대 13.07%로 축소되었다. 기본적으로 원본 씬을 변화시키지 않는 범위에서 최적화가 진행되어야 한다. 하지만 씬의 사이즈가 최적화 이후에도 원하는 사이즈보다 클 경우 부득이 씬의 정보를 변화시키는 최적화 작업을 진행해야 한다. 따라서 원본 씬과의 차이는 BIFS 커맨드를 사용하여 실시간으로 씬을 업데이트 시켜야 한다. BIFS 커맨드가 또 다른 지연을 발생시킬 수 있는 경우를 감안하여 효율적인 TS 패킷화를 고려한 대화형 데이터의 커맨드 단위화에 대한 연구가 진행 중이다.

참고 문헌

[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] Yong He, Ahmad, I., Liou, M.L "Real-time Interactive MPEG-4 System Encoder using a cluster of workstations," IEEE Transactions on Multimedia, Volume 1, Issue 2, pp.217-233, June 1999.

[3] Zhaohui Cai, K.R. Subramanian, Aidong Men, Yong Ji, "A RISC Implementation of MPEG-2 TS Packetization", High Performance Computing in the Asia-Pacific Region, Proceedings. The Fourth International Conference/Exhibition on Volume 2, pp.688 - 691, May 2000.

[4] K.Cha, S. Kim "MPEG-4 STUDIO: An Object-Based Authoring System for MPEG-4 Contents," Multimedia Tools and Applications, Volume 25, Issue 1, pp.111 - 131, January 2005.