

# 모바일 환경을 위한 미디어 형식 변환 모델 연구<sup>1</sup>

한병준\*<sup>0</sup> 윤정현\*\* 이한규\*\* 황인준\*  
\*고려대학교 대학원 전기전자전자공학부  
\*\*한국전자통신연구원(ETRI)

hbj1147@mil.korea.ac.kr<sup>0</sup>, { sine, hkl }@etri.re.kr, ehwang04@korea.ac.kr

## A Study on the Media Transmodality Model for Mobile Environment

Byeongjun Han\*<sup>0</sup>, Jeonghyun Yoon\*\*, Hankyu Lee\*\*, Eenjun Hwang\*  
\*Department of Electronics and Computer Engineering, Korea University  
\*\* Electronics and Telecommunications Research Institute

### 요 약

임베디드 기기와 무선 네트워크 기술이 발전함에 따라 PDA나 휴대폰과 같은 모바일 기기 상에서도 멀티미디어 콘텐츠를 소비하는 것이 가능하게 되었다. 그러나 모바일 기기 특성상 무선 네트워크를 통한 콘텐츠 소비 환경은 유동성이 강하다. 따라서 모바일 환경에서 멀티미디어 콘텐츠를 효과적으로 재공하기 위해서는 상황에 따른 가용적인 모델이 필요하다. 미디어 형식 변환(Media Transmoding)은 네트워크, 하드웨어, 스트리밍 서버 등의 상황에 따라 능동적으로 콘텐츠의 형식을 변환하여 전송하는 기술이다. 본 논문에서는 모바일 환경에서의 미디어 형식 변환 모델을 제안하고, 자체 개발한 미디어 형식 변환 시나리오를 사용하여 TV-AnyTime Phase 2 Part 3와 SOAP과 결합하여 설계 및 구현하였다.

### 1. 서 론

최근 들어, 하드웨어와 소프트웨어의 발전에 의해 가용한 콘텐츠의 양이 폭발적으로 증가하였다. 또한, 사용자는 원하는 콘텐츠를 자신만의 환경에서 보고자 하는 욕구가 증대되었다. 이에, PDA나 휴대폰과 같은 다양한 무선 모바일 환경에서 무선단말기를 기반으로 하는 멀티미디어 콘텐츠 소비 시장이 활성화 되었다.

그러나 모바일 환경에서의 멀티미디어 콘텐츠 활용은 무선 네트워크의 특성상, 사용자가 현재 속해있는 환경이나 위치, 지형, 그리고 사용자가 가지고 있는 기기에 따라 재생할 수 있는 미디어 형태에 한계가 있다. 따라서 상황에 따라 적절한 미디어 형태를 결정하고 이를 재생할 수 있는 모델이 필요하다.

본 논문에서는 모바일 환경에서 다양한 제약조건에 따라 미디어를 효과적으로 재생할 수 있는 미디어 형식 변환 모델(Media Transmoding Model)을 제안한다. 일반적인 모바일 기기는 컴퓨팅 능력이 그리 뛰어나지 않기 때문에, 기존 콘텐츠의 비디오, 오디오, 이미지, 텍스트와 같은 단순화된 모달리티(modality)의 표현으로 다양한 대역폭 및 기기 상황, 유저의 기호에 대응할 수 있다.

또, TV-Anytime 중 TV-Anytime Phase 2 중 Part 3 규격에 명시된 메타데이터 형식을 사용하여 모바일 기기, 사용자 환경, 그리고 사용자 기호 등을 XML 스키마로 기술함으로써, 서버-클라이언트 모델에 있어서 SOAP과 같은 XML 기반 프로토콜을 사용할 수 있게 되어있다.

논문의 이후 부분은 다음과 같이 구성되어 있다. 2장에서는 미디어 형식 변환(Transmoding)과 관련한 다른 연구 사례를 보

이고 있으며, 3장에서는 형식 변환 모델의 전반을 설명한다. 4장에서는 해당 프레임워크의 실제 구현을 설명하며, 5장에서는 현재까지 구현 내용을 정리하고 향후 연구 과제를 전망하도록 한다.

### 2. 관련 연구

사용자의 환경에 적합한 미디어 형식 변환을 위한 연구는 최근 스트리밍 환경이 유/무선, 모바일, PMP 등으로 다변화되고 유비쿼터스 환경에 대한 관심이 늘어나면서 활발히 이뤄지고 있다. 사용자가 언제 어디서나 어떤 기기를 가지고도 미디어를 향유하는 것을 목적으로 하는 UMA(Universal Multimedia Access)의 개발을 위한 표준 및 연구가 많다[1][2][3].

미디어 형식 변환(Transmoding/Transcoding)에 대한 연구도 많다. 대역폭 감쇄, 형식 전환, 모달리티(Modality) 제거 등과 같은 연구가 이뤄져 왔다[7][8][9][10]. 또한, 모달리티 간의 전환을 줄더 능동적으로 하기 위해 모달리티 간의 값을 설정한 OCV(overlapped content value) 모델 등도 있다[11].

미디어 데이터의 메타데이터 정보를 표준화하기 위한 연구 및 표준화 작업 또한 활발히 이뤄졌다. MPEG-7 표준[4][5][6]은 멀티미디어 데이터와 사용자 정보를 포괄하는 중요한 메타데이터 프레임워크를 제공하였다. 그러나 MPEG-7 표준이 주로 미디어 검색을 위한 메타데이터를 정의하고 있다면, 방송 표준인 MPEG-21[1]은 이를 확장하여 MPEG-21 Part 7 " Digital Item Adaptation(DIA)[3]" 과 같이 디지털 미디어의 포괄적인 정의를 제공하였다.

<sup>1</sup> 본 연구는 2005년도 정보통신부 및 정보통신연구진흥원의 대학 IT 연구센터 지원사업과 한국전자통신연구원(ETRI) 모바일 소비환경을 고려한 미디어 형식변환 모듈 개발의 연구용역 지원으로 수행하였습니다.

3. 미디어 형식 변환 모델

본 논문에서 제안하는 모델의 서버-클라이언트 프로그램은 스트리밍 서버로부터 미디어 형식 변환 기능을 제공하고 있다. 또, 사용자 측의 정보를 서버에 전달하기 위한 메타데이터로써 TV-Anytime 중 TV-Anytime Phase 2에 정의된 메타데이터를 활용하고 있다. 이와 관련한 전체 구조는 그림 1에 나타나 있으며, 각 부분에 대한 세부 설명은 다음과 같다.

3.1. 미디어 형식 변환(Media Transcoding)

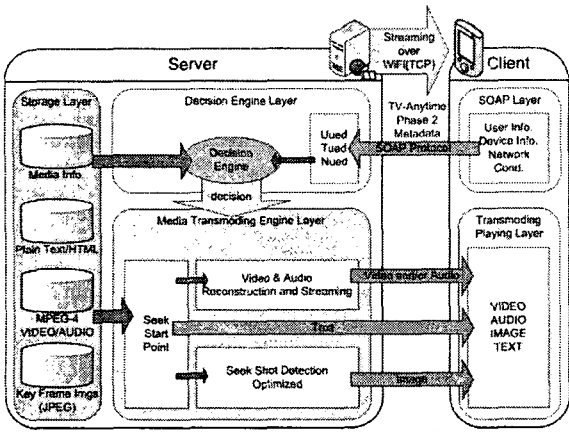


그림 1 전체 미디어 형식 변환 모델

미디어 형식 변환(Media Transcoding)은 사용자의 기호나 환경에 맞게 비디오, 오디오, 이미지, 텍스트 자막과 같은 형태로 콘텐츠를 가공하여 서버에서 클라이언트로 전송, 표시하는 과정을 말한다. 이 과정에서의 미디어 형식을 모달리티(Modality)라 한다.

제안하는 미디어 형식 변환 모델을 설명하면 다음과 같다. [그림 1] 중 데이터 계층(Storage Layer)에는 원래의 동영상, 원 동영상으로부터 추출한 이미지, 그리고 텍스트 자막 데이터가 있으며, 이미지 및 텍스트와의 동기화 정보는 미디어 정보의 형태로 담겨있다. 미디어 형식 변환 엔진 계층(Media Transcoding Engine Layer)은 클라이언트로부터 원하는 모달리티, 재생 시작 정보를 가지고 데이터 계층으로부터 적절한 데이터를 받아 처리하여 클라이언트에게 전송한다.

3.2. 전체 모델의 기술

전체 모델은 계층 간의 연계로 이루어져 있다. [그림 1]에서 사용자는 미디어 공급자로부터 제공받은 URL 주소를 바탕으로 미디어 형식 변환 서버에 질의한다. 이때 주기적으로 기기에 설정해놓은 사용자 정보, 기기 정보, 그리고 네트워크 상황을 포함한 환경 정보를 TV-Anytime Phase 2 규격 메타데이터로 전송한다. 이후, 스트리밍 서버 측의 미디어 결정 계층(Decision Engine Layer)은 동영상의 정보와 사용자로부터의 정보를 바탕으로 모달리티를 결정한다. 결정된 모달리티는 클라이언트에도 전달되며, 이후 사용자는 상황에 따라 적절한 모달리티를 지정 받아 미디어 콘텐츠를 감상할 수 있다.

3.3. 미디어 형식 변환 시나리오

[그림 2]은 미디어 형식 변환 요청 발생시 새로운 모달리티의

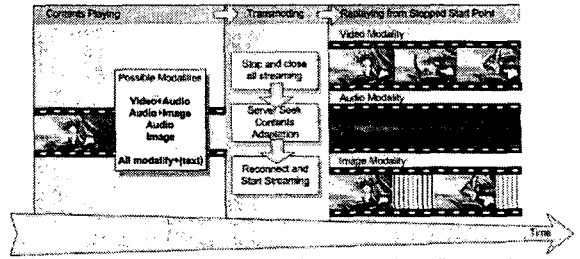


그림 2 미디어 형식 변환시 새 모달리티 재생 과정

재생 과정을 보여주고 있다. 우선 클라이언트로부터 전달된 메타데이터를 적절히 판별하여 새로운 모달리티를 도출한다. 이후, 새로운 모달리티를 사용하여 스트리밍 서버에 접속한다. 이때 클라이언트가 재생하던 지점으로부터 재생이 재개되어야 하므로 클라이언트는 서버에 재생 지점에 대한 인자를 전달한다. 이후 적절한 재생 지점을 탐색하여 재생 요구 지점에 최적으로 가까운 콘텐츠를 계산 및 추출 전송한다.

3.4. TV-Anytime Phase 2 Metadata [13] 표준의 적용

TV-Anytime Phase 2는 네트워크가 구비된 환경에서 단일 시스템 간의 콘텐츠 공유 및 재배포, Targeting 등의 서비스를 제공하는 맞춤형 양방향 방송기술의 표준이다. 특히 TV-Anytime Phase 2의 표준 규격 구성 중 Part 3 메타데이터 규격은 XML 기반 프로토콜인 SOAP을 사용하여 메타데이터를 전송할 것을 권고하고 있으며, 규격 중 Subpart 3 메타데이터 확장 [13]은 Targeting Information을 XML 스키마로 정의하고 있다.

본 논문의 목적은 모바일 상황 정보에 대한 미디어 형식 변환 시나리오의 적용이므로, 다양한 스키마 형식 가운데 상황을 가장 잘 나타내는 User, Terminal, Network Information을 적용하여 SOAP 기반으로 개발을 진행하였다.

4. 시스템의 설계와 구현

본 논문에서는 [그림 1]의 모델을 바탕으로 스트리밍 서버 Livsend(Win)와 PDA 미디어 형식 뷰어인 Displayer2를 개발하였다. 동영상 모달리티 재생 기본 코덱은 MPEG-4이지만, WMV(Windows Media Video)가 허용하는 코덱 형식이면 모두 재생할 수 있도록 설계되어 있다.

4.1. 전체 개발 환경

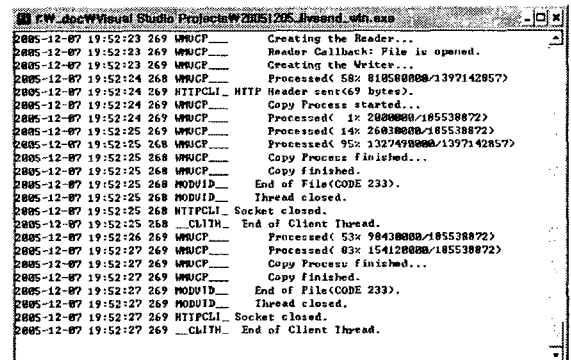


그림 3 Livsend(Win)의 작동 화면

[그림 3]는 스트리밍 서버 Livsend(Win)의 작동 화면을 나타내고 있다. Livsend(Win)은 펜티엄 IV 3.0GHz, 512MB RAM, 40GB HDD, 100Mbps의 환경에서 개발 및 테스트되었으며, 구현 언어로는 Windows 플랫폼 기반의 C++을 사용하였다. 또, Windows 기반 스트리밍 핵심 라이브러인 COM 기반의 Windows Media Format을 사용하였다.

[그림 4]은 PDA용 미디어 형식 변환 플레이어인 Display2의 작동 화면이다. eMbedded C++ 4.0을 이용하여 Pocket Windows 2003 플랫폼 기반으로 개발되었으며, 테스트는 HP IPAQ HX4700 모델(625MHz, 64MB ROM, 480 x 640)에서 이뤄졌다.

서버와 클라이언트 모두 기반 프로토콜로 TCP-HTTP를 선택한 것은, UDP 기반에 비해 일반적인 스트리밍 환경과 호환성이 높으며 QoS를 위한 별도의 제어 규칙을 정의하지 않아도 되기 때문이다[12].

#### 4.2. 구현 및 테스트 결과

테스트에 사용된 미디어 콘텐츠는 다음과 같다. 비디오는 MPEG4 코덱으로 320x240@30FPS에 평균 300Kbps로 인코딩하였으며, 오디오는 96Kbps MP3 코덱을 사용하였다. 그리고 파일 패키지 형식으로는 WMV(Windows Media Video)를 사용하였다. 테스트에 사용된 동영상은 YTN 뉴스, 매트릭스 PV(각각 약 30sec.), 각종 TV CF(각각 약 15sec.), 온라인 게임 PV(약 5min.), 일본 애니메이션(약 20min.) 등이었다.

무선 네트워크 환경으로는 고려대학교 자연계 캠퍼스의 WiFi 환경이 적용되었다. 서버는 일반적인 100Mbps 유선 네트워크에 연결하고, 무선 서비스의 대상이 되는 PDA가 캠퍼스의 이곳 저곳을 방문하며 모달리티 변경 및 음영 지역에서의 퍼포먼스를 테스트하였다.

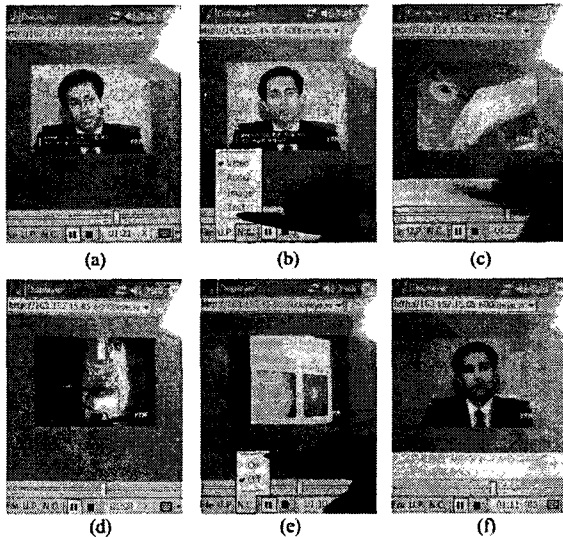


그림 4 Display2로 모달리티 변경 화면

[그림 4]의 (a)(b)(c)는 Decision Engine에 의한 Feedback을 제거한 상황이다. (a)는 일반적인 동영상 스트리밍 되고 있다. (b)에서는 사용자가 텍스트 모달리티를 선택하고 있으며, 이러한 모달리티를 선택하면 (c)와 같이 텍스트 정보를 받을 수 있다. 반면, (d)(e)(f)는 Decision Engine을 작동시키고 50~

100Kbps 정도로 대역폭을 낮췄을 때의 작동 화면이다. (d)는 Decision Engine을 작동시키기 이전을 보여주고 있다. 비디오+오디오 모달리티로 작동하고 있었으나, (e)에서 Decision Engine을 작동시킨 후, 대역폭이 낮아지자 서버와의 피드백을 통해 (f)에서 이미지+텍스트 모달리티로 변환된 것을 보이고 있다.

#### 5. 결론과 향후 연구 과제

본 논문에서는 모바일 환경에서 미디어 형식 변환(Media Transmoding)을 지원하는 서버-클라이언트 모델을 제안하고 이를 설계 및 구현하였다. 특히 제안된 모델은 맞춤형 방송 규격인 TV-Anytime에 기반한 메타데이터를 사용하여 모바일 환경에 따른 능동적인 형식 변환 엔진을 제공하고 있다.

향후 와이브로(WiBro)와 같은 광역 휴대 인터넷이 보급되면 무선 환경을 통한 미디어 콘텐츠의 소비가 늘 것으로 예상된다. 특히, 주변 유동 환경에 따라 최적의 미디어 형식 변환을 제공하는 등의 강력한 미디어 형태 변환 기술의 연구가 요구될 것이다.

또한, 개인 스트리밍 서버와 방송 채널의 경계가 점점 허물어질 것이다. 따라서 멀티미디어 콘텐츠를 편리하게 볼 수 있는 새로운 인터페이스의 정의 또한 필요할 것이다.

#### 참고 문헌

- [1] MPEG Requirements Group, "MPEG-21 Multimedia framework, Part 1: Vision technologies and strategy? Proposed Draft Technical Report", 2<sup>nd</sup> Edition, Doc. ISO/MPEG N6269, MPEG Waikaloa Meeting, USA, December 2003.
- [2] MPEG MDS Group, "Multimedia framework Part 7: Digital item adaptation, FinalDraft International Standard," Doc. ISO/MPEG, N6168, MPEG Waikaloa Meeting, USA, December 2003.
- [3] Anthony Vetro, Christian Timmer, "Digital Item Adaptation: Overview of Standardization and Research Activities," IEEE Trans. Multimedia, Vol. 7, No 3, June 2005.
- [4] ISO/IEC JTC1 SC29/WG11 N3465, "MPEG-7 Multimedia Description Schemes WD," Beijing, 2000.
- [5] R. S. Ramanujan, J. A. Newhouse, M. N. Kaddoura, A. Ahamad, E. R. Chartier, and K. J. Thurber, "Adaptive Streaming of MPEG Video over IP Networks," Pro. Of the 22<sup>nd</sup> IEEE Con. On Computer Networks (LCN 97), November 1997.
- [6] J. Magalhaes, F. Pereira, "Using MPEG standards for multimedia customization," Signal Processing: Image Communication. Vol. 19, pp. 437-456, 2004.
- [7] W. Y. Ma, I. Bedner, G. Chang, A. Kuchinsky, and H.J. Zhang, "A framework for adaptive content delivery in heterogeneous network environments," SPIE Multimedia Computing and Networking, San Jose, California, USA, 2000.
- [8] C.G.M. Snoek, M. Worring, "Multimodal video indexing: A review of the state-of-the-art," Multimedia Tools and Applications, 2004.
- [9] Sumi Halal, Latha Sampath, Keven Birkett and Joachim Hammer, "Adaptive delivery of video data over wireless and mobile environments," Wireless Communications and Mobile Computing, Vol. 3, pp. 23-36, 2003.
- [10] W. Y. Han, O. C. Kwon, J. H. Park, J-H Kang, "A Gateway and Framework for Telematics Systems Independent on Mobile Network," ETRI Journal, Vol. 27, No 1, Feb. 2005.
- [11] T. C. Thang, Y. J. Jung, Y. M. Ro, "Modality conversion for QoS management in universal multimedia access," IEEE Proc.-Vis. Image Signal Process, Vol. 152, No 3, June 2005.
- [12] B. Wang, J. Kurose, P. Shenoy, D. Towsley, "Multimedia Streaming Via TCP: An Analytic Performance Study", ACM Multimedia 2004, pp. 908-915, October 2004.
- [13] ETSI TS 102 822-3-3: "TV-Anytime Part 3: Metadata, Sub-part 3: Phase 2 - Extended Metadata Schema", Final Edition, Doc., EBU/ETSI JTC Broadcast, January, 2006.