

UMA를 위한 유틸리티 기반 MPEG-21 비디오 적응

김재곤, 강경옥, 김진웅, 김형명*

한국전자통신연구원 방송미디어연구부, *KAIST 전자전산학과

전화: 042-860-4980/팩스: 042-860-5479

Utility-Based MPEG-21 Video Adaptation for Universal Multimedia Access

Jae-Gon Kim, Kyeongok Kang, Jinwoong Kim, and Hyung-Myung Kim*

Dept. of Broadcasting Media Technology, ETRI, Dept. of EECS, KAIST

jgkim@etri.re.kr

Abstract

Video adaptation in response to dynamic resource conditions and user preferences is required as a key technology to enable universal multimedia access (UMA) through heterogeneous networks by a multitude of devices in a seamless way. Although many adaptation techniques exist, selections of appropriate adaptations among multiple choices are often ad hoc. To provide a systematic solution, we present a general conceptual framework to model video entity, adaptation, resource, utility, and relations among them. It allows for formulation of various adaptation problems as resource-constrained utility maximization. We apply the framework to a practical case of dynamic bit rate adaptation. Furthermore, we present a description tool, which has been accepted as a part of the MPEG-21 Digital Item Adaptation (DIA), along with a brief overview of the related descriptors to support terminal and network quality of service (QoS).

I. 서론¹

최근 이종망과 다양한 사용자 단말(TV, PC, PDA, cellular phone 등)을 통하여 언제 어디서나 끊김없이 멀티미디어 콘텐츠를 접근/소비할 수 있는 UMA(Universal Multimedia Access) 멀티미디어 프레임워크가 대두되고 있다[1]. UMA 프레임워크에서 이종의 동적 리소스 제약(resource constraints)과 사용자 선호도 등에 대응하기 위한 콘텐츠 적응(adaptation)이 핵심요소 기술로 요구된다. 특히 UMA와 같이 다양한

망과 사용자 및 단말로 구성된 멀티미디어 프레임워크에서 상호연동 가능한 단-대-단 적응 기술이 요구된다. MPEG에서는 이러한 멀티미디어 응용 서비스를 위한 기술로 “MPEG-21: 멀티미디어 프레임워크[1]”의 표준화를 진행하고 있다. 이 중 파트 7: 디지털 아이템 적응(DIA: Digital Item Adaptation)[2]에서 네트워크 및 단말, 사용자 선호도 등의 사용자환경(Usage Environment), 그리고 미디어 리소스 및 메타데이터 적응을 위한 정보를 기술하기 위한 기술자(descriptor)를 표준화하고 있다. 본 논문은 MPEG-21 DIA의 터미널 및 네트워크 QoS를 위한 기술자로 제안, 채택된 유틸리티 기반의 적응 기법 및 관련 기술자들을 소개한다.

기존의 비디오 적응 기법[4]으로는 재양자화(requantization)[5], CD(DCT Coefficient Dropping)[6], 공간해상도 변형 등의 공간영역 적응과 FD(Frame Dropping)[7] 등의 시간영역 적응이 있다. 그러나 기존의 적응 기법들은 주로 선택된 적응을 최적화하는데 집중되었으며, UMA 환경에서의 다양한 동적 제약을 만족시키는 최적의 적응기(adaptation operator)를 선택하기 위한 체계적인 기법을 제시하지 못하였다. 본 논문에서는 기존의 율-왜곡 (Rate-Distortion) 모델을 확장한 유틸리티 기반의 적응 프레임워크를 제시하고자 한다. 본 프레임워크에서는 주어진 적응 문제를 적용하고자 하는 ‘적응 (Adaptation)’, 가용한 ‘리소스(Resource)’, 적응결과 ‘유틸리티(Utility)’, 즉 ‘A-R-U’의 관계를 ‘유틸리티 함수(utility function)’로 모형화함으로써 최적의 적응기를 선택 가능하도록 한다. 본 프레임워크는 율-왜곡 모델에서 다루지 못한 다양한 종류의 리소스, 유틸리티 및 적응을 포함한다.

본 프레임워크의 타당성 검증을 위한 실례로 시변 가용 대역폭에 적응하기 위한 유틸리티 기반의 율적응(rate adaptation)으로 FD와 CD를 결합한 ‘FD-CD’

¹ 본 연구는 정보통신부 과제의 일부로서 수행되었습니다.

변환(transcoding)을 제시한다.

II. 유틸리티 기반 적응 프레임워크

유틸리티 기반 적응 프레임워크는 다음 3 단계를 포함한다: 1) 비디오 엔티티(entity) 및 적응하고자 하는 적응에 대한 정의, 2) 유틸리티 함수를 이용한 정의된 모든 적응기에 대한 리소스와 유틸리티의 모형화, 3) 유틸리티 함수 기반의 최적 적응기 선택. 본 프레임워크의 핵심은 적응 문제를 A-R-U로 모형화하는 것이며, 이들 변수들은 다차원의 공간(space)으로 일반화된다.

먼저 비디오 엔티티는 적응 단위를 정의하는 것으로 이 엔티티 마다 유틸리티 함수를 생성한다. 리소스-유틸리티의 관계가 콘텐츠 특성에 의존적이므로 유사한 특성을 갖는 비디오 세그먼트로 정의한다. 적응공간(Adaptation Space)의 각 차원은 특정 적응방법을 나타내고 각 적응방법에서 유한 개의 가능한 적응기가 정의된다. 예를들면 FD-CD의 경우 적응공간은 FD와 CD의 2 차원 공간으로 구성되고 각 차원은 몇 가지의 가능한 적응기(예, FD: all B-frames dropping, all B&P-frames dropping, CD: 10%, 20%, ..., 50% rate-reduction, 등)로 색인된다. 따라서 적응공간은 각 차원에서 정의된 적응기를 조합한 FD-CD의 모든 가능한 적응기의 분포를 포함한다. 리소스공간(Resource Space)은 주어진 대역폭, 단말화면 크기, 메모리 크기 등 동시에 만족시켜야 할 제약요를 표현한다. 유틸리티 공간(Utility Space)은 PSNR, 충실도(fidelity), MOS(mean opinion scale) 등 다 수의 객관적/주관적 측정에 의한 화질을 표현한다.

유틸리티 함수는 적응공간에 정의된 모든 적응기에 대한 유틸리티-리소스의 값을 유틸리티- 리소스 평면에 나타냄으로써 A-R-U 공간의 매핑관계를 효율적으로 표현한다. 이 유틸리티 함수는 MPEG-21 DIA의 기술자로 기술되어 해당 비디오 비트열과 함께 전송되어 네트워크 노드에서의 체계적인 실시간 최적 적응을 가능하게 한다.

그림 1 은 유틸리티 기반의 적응 프레임워크의 전체적인 구조를 나타낸 것으로, 저장 또는 라이브로 입력되는 비디오 비트열에 대하여 유틸리티 함수를 생성하는 서버, 전송된 비트열을 다양한 사용자 단말 및 접근네트워크의 제약에 대응하여 유틸리티 함수를 이용하여 실시간 적응을 하는 적응엔진으로 구성된다.

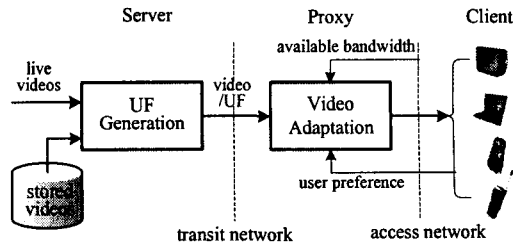


그림 1. 유틸리티 기반의 비디오 적응 프레임워크의 구성도

저장된 비디오의 경우 유틸리티 함수는 정의된 모든 적응기에 대한 리소스-유틸리티 값을 미리 계산하여 생성할 수 있는 반면, 라이브 비디오의 경우 많은 계산량을 요하는 계산에 의한 실시간 생성이 불가능하며 콘텐츠 기반의 예측의한 접근법[3] 사용할 수 있다.

III. 유틸리티 기반 FD-CD

다양한 단말을 수용하기 위한 이중망의 통합과 무선 채널의 가변 채널상태 등의 영향으로 인한 시변 가용 대역폭에 대응하기 위한 율적응은 UMA 환경에서의 전형적인 비디오 적응 예이다. 이 경우 비디오 적응은 네트워크 노드에 위치한 적응엔진에서 실시간으로 수행되어야 하고 넓은 범위의 대역폭 변화를 수용하여야 한다. 본 논문에서는 이러한 요구사항을 고려하여 UMA를 위한 율적응 기법으로 FD와 CD를 결합한 FD-CD를 제시한다.

FD와 CD는 각각 일반적인 공간 및 시간적응 기법으로 비트열에서 특정 프레임 및 DCT 계수에 해당하는 부분을 제거함으로써 율을 감소시키는 방법으로 실시간 응용에 적합하다. FD의 경우 최소 적응 단위가 프레임이므로 몇 가지의 제한적인 율 적응만이 가능한데 비해 CD는 제거되는 계수를 조정함으로써 주어진 대역폭을 더 정확하게 만족시킬 수 있다. 따라서 두 가지 기법의 결합을 통하여 넓은 범위의 대역폭 변화에 따른 목표율(target rate)을 충분히 정확하게 얻을 수 있다. 또한 콘텐츠 내용에 따라서 공간 및 시간적응의 정도를 조정함으로써 시공간 화질간의 최적 조정이 가능하다.

FD-CD의 A-R-U 공간은 다음과 같이 정의된다. FD와 CD로 구성된 2 차원의 적응공간, $A = \{a_i = (f_i, c_i)\}$ 로 정의되고, 여기서 $a_i = (f_i, c_i)$ 는 FD-CD의 적응기로 입력 엔티티, e 에 적용된다. f_i 와 c_i 는 각각 FD와 CD에서 정의된 적응기를 나타낸다.

리소스공간과 유틸리티공간은 각각 적용된 엔티티, $e'_i = \mathbf{a}_i(e)$ 의 율과 PSNR을 측정하는 $rate(e'_i)$ 과 $PSNR(e'_i)$ 로 정의된다.

예로 그림 2의 유틸리티 함수는 GOP(N=15, M=3)의 구조를 갖는 입력 비디오에 대하여 $f_i = \{no, B, BP\}$, $c_i = \{0, 0.1, 0.2, 0.3, 0.4, 0.5\}$ 의 조합에 의한 18개의 FD-CD 적용기(그림의 각 포인트에 해당)에 대한 율과 PSNR 값을 나타낸 것이다. 여기서 f_i 의 각 요소는 no-dropping, all-B-frames -dropping, all-B&P-frames-dropping을, c_i 는 CD에 의한 율감소 비율을 나타낸다.

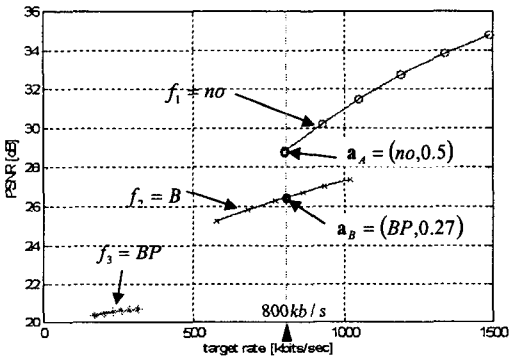


그림 2. FD-CD 적용에 대한 유틸리티 함수 ("coastguard" coded at 1.5Mbps in MPEG-4)

FD와 달리 CD의 경우 인접한 서로 다른 CD간의 보간이 가능하기 때문에 그림 2와 같이 동일한 FD와 서로 다른 CD 적용을 갖는 적용기들은 한 곡선으로 연결될 수 있다. 이러한 유틸리티 함수가 제공될 경우, 그림 2에서와 같이 목표 율이 800kp/s로 주어진 경우 이를 만족하는 적용기는 $\mathbf{a}_A = (no, 0.5)$ 와 $\mathbf{a}_B = (B, 0.27)$ 가 존재하고 이중에서 \mathbf{a}_A 가 PSNR 측면에서 최적의 적용기임을 알 수 있다.

IV. 유틸리티 함수 기반 기술자

본 논문에서 기술한 적용-리소스-유틸리티의 정보를 기술하는 유틸리티 함수 기반의 기술자(descriptor)[8]를 MPEG-21 DIA의 터미널 및 네트워크 QoS를 위한 미디어리소스 적용 툴로 제안하여 현재 *AdaptationQoS*[2],[9]의 일부로 채택되어 최종위원회초안 단계의 표준화 작업이 진행 중이다.

그림 3은 *AdaptationQoS* 기술자의 구성도이다. *AdaptationQoS*는 다수의 *AdaptationQoSModule*과 각

모듈의 입출력을 기술하는 *IOPin*으로 구성된다. 모듈은 하나의 기술 단위로 기술하고자 하는 데이터의 특성에 따라서 *LookUpTable*, *StackFunction*, 또는 *UtilityFunction*으로 상속 구현된다[9]. 본 논문에서 제안한 *UtilityFunction*은 그림 2와 같은 유틸리티 함수를 리소스 축을 유한개의 포인트로 표본화한 *Constraint*와 그 각 포인트에서의 리소스를 만족시키는 적용기를 기술하는 *AdaptationOperator*, 각 적용기에 의한 결과 유틸리티 값을 기술하는 *UtilityValue*로 구성된다. *AdaptationOperator*는 확장성을 위하여 *ClassificationScheme*으로 정의하며 현재 *Frame Dropping*, *Coefficient Dropping*, *MPEG-4 FGS*, *Spatial Size Reduction*, *Requantization* 및 *Audio FGS* 등의 적용 기법과 이들 기법의 적용기를 기술하기 위한 파라미터의 정의를 포함하고 있다.

*UtilityValue*의 경우 CD와 같이 기술되는 적용기에 의한 유틸리티가 구현 방법에 의존적인 경우 유틸리티 값 대신에 주어진 리소스제약을 만족하는 다수의 적용기들간의 유틸리티 값에 따른 순서를 *UtilityRank*로 기술함으로써 기술(description)의 모호성의 문제를 해결한다[8].

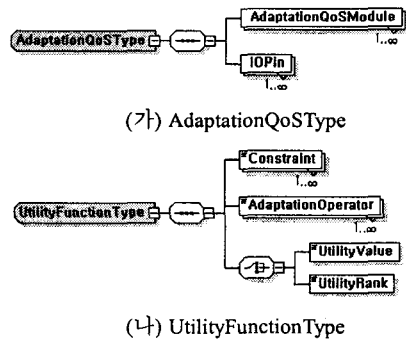


그림 3. *AdaptationQoS* 기술자의 구성도

그림 4는 MPEG-21 DIA의 미디어리소스 적용 툴에 기반한 적용 구성도이다. 압축 비트열의 상위구조를 기술하는 비트열신택스기술(BSD) 툴인 *gBSD/BSDL*는 *AdaptationQoS*에 기반하여 선정된 적용기에 따라 입력 미디어리소스 비트열을 압축 포맷 독립적인 효율적인 방법으로 적용한다. 이들 툴의 상호연동을 위해서는 각 툴에서 적용을 기술하는 파라미터간의 연결이 필요하며 이를 위한 *BSDLink*[9] 툴에 대한 표준화 작업이 진행 중이다. 그림 4의 적용엔진을 비롯한 각 구성모듈의 기능구현은 표준화에 포함되지 않는다.

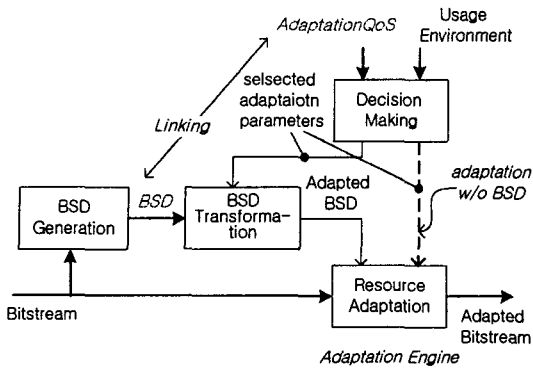


그림 4. BSDLink 틀을 활용한 리소스 적응 구성도

V. 실험결과

본 논문에서는 유틸리티 기반의 율적용 FD-CD 프로토타입 틀을 구현하고 MPEG-4 시퀀스를 사용하여 성능 평가를 하였다. 본 프로토타입은 파서(parser), 적응엔진, 그리고 사용자인터페이스 (GUI)로 구성된다. 파서는 각 비디오 엔티티에 해당하는 AdaptationQoS 기술자를 파싱하고 유틸리티 함수를 재구성하여, 주어진 목표 율을 만족하는 적응기를 선택해서 적응엔진으로 전달한다. 적응엔진은 MPEG-4 비트열을 엔티티 단위로 선택된 FD-CD 적응을 수행하여 적응된 비트열을 출력한다. 그림 5의 GUI는 적응과정을 모니터링하기 위한 것으로, 비디오 창에서는 적응된 비디오의 화질을 확인하고 상단 오른쪽 창은 재구성된 유틸리티 함수와 하단의 시변 대역폭에 적응하기 위한 최적 적응기의 선택과정을 도시한다. 사용자선호도 입력창은 사용자가 PSNR에 의한 최적 적응기의 자동 선택 대신에 공간 또는 시간화질의 사용자 선호도에 따라서 적응기를 선택할 수 있도록 한다.

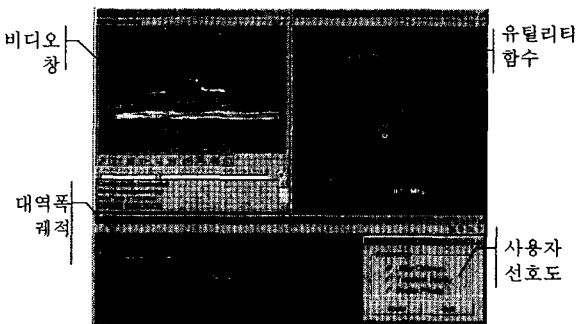


그림 5. 유틸리티 기반 비디오 적응 틀의 GUI

본 틀을 이용한 실험에서 유틸리티 함수를 이용하지 않는 기존의 방법(실험 데이터 셋의 사전정보로부터 율에서의 가장 많이 사용된 최적 적응기를 콘텐츠에 무관하게 사용하는 방법으로 가정)과 과 비교하여 0.83 dB의 PSNR 이득과 주관적인 화질 개선을 확인 하였다.

VI. 결론

본 논문에서는 UMA를 위한 유틸리티 기반의 비디오 적응 프레임워크를 제시하였다. 본 프레임워크는 서버에서 적응-리소스-유틸리티의 관계를 유틸리티 함수를 이용하여 모형화하고 미디어리소스와 함께 전송함으로써 네트워크 노드에서의 실시간으로 체계적인 최적적용이 가능하도록 하였다. 본 프레임워크의 타당성 검증 예로 유틸리티 기반의 FD-CD 율적용을 제시하였다.

또한 MPEG-21 DIA에 제안/채택된 유틸리티 함수를 기술하는 기술자 및 관련 미디어 리소스 적응 틀들을 기술하였다.

참고문헌

- [1] J. Bormans, J. Gelissen, and A. Perkis, "MPEG-21: The 21st century multimedia framework," *IEEE Signal Processing Mag.*, vol. 20, pp. 53-62, Mar. 2003.
- [2] Study of ISO/IEC 21000-7 CD - Part 7: Digital Item Adaptation, ISO/IEC JTC1/SC29/WG11/N5612, Pattaya, Thailand, Mar. 2003.
- [3] Y. Wang, J.-G. Kim, and S.-F. Chang, "Content-based utility function prediction for real-time video transcoding," in *Proc. IEEE ICIP2003*, Sep. 2003, Barcelona, Spain.
- [4] A. Vetro, C. Christopoulos, and H. Sun, "Video transcoding architectures and techniques: An overview", *IEEE Signal Processing Mag.*, vol. 20, pp. 18-29, Mar. 2003.
- [5] Y. Nakajima, H. Hori, and T. Kanoh, "Rate conversion of MPEG coded video by requantization process," in *Proc. IEEE ICIP'95*, Washington, DC, 1995, pp. 408-411.
- [6] A. Eleftheriadis, "Dynamic rate shaping of compressed digital video," Ph. D. dissertation, Dept., Elec. Eng., Columbia Univ., New York, June 1995.
- [7] K. T. Fung, Y. L. Chan, and W. C. Siu, "New architecture for dynamic frame-skipping transcoder," *IEEE Trans. Image Processing*, vol. 11, pp. 886-900, Aug. 2002.
- [8] J.-G. Kim, Y. Wang, S.-F. Chang, K. Kang, J. Kim, "Description of utility function based optimum transcoding," ISO/IEC JTC1/SC29/WG11/M8319, Fairfax May 2002.
- [9] MPEG-21 Digital Item Adaptation Am (v5.0), ISO/IEC JTC1/SC29/WG11/N5613, Pattaya, Thailand, Mar. 2003.