

QoS 적응을 위한 MPEG-2 비디오의 스케일러블 미디어 객체 프레임워크에 관한 연구

김 형 철[†] · 정 찬 근^{††}

요 약

본 논문에서는 이질적인 통신망 환경에서 다양한 QoS 요구에 적용할 수 있는 MPEG-2 비디오의 스케일러블 미디어 객체 프레임워크를 제안한다. 하나의 MPEG-2 비디오는 시간적인 스케일링과 충실도 스케일링 과정을 통하여 상호 종속성이 있는 스케일러블 미디어 객체의 집합으로 변환된다. 미디어 서비스 요청에 대하여 스케일러블 미디어 객체의 전체 혹은 일부를 전송함으로써 다양하게 요구된 QoS에 따른 전송 비트율을 만족시킬 수 있다. 또한, 다양한 대역폭의 링크가 개입된 멀티포인트 채널에 대해서도 각 QoS를 만족시킬 수 있는 기반을 제공한다. 전송된 데이터 양과 신호왜곡 특성을 분석을 통하여 스케일러블 미디어 객체가 폭넓은 QoS 요구에 적용할 수 있음을 보였다.

Scalable Media Object Framework of MPEG-2 Video for QoS Adaptation

Hyung-Chul Kim[†] · Chan-Gun Jeong^{††}

ABSTRACT

In this paper, we propose a new framework for scalable media object to adapt the various QoS requirements in the heterogeneous networking environment. An MPEG-2 video is split into a set of scalable media objects in terms of temporal scaling and fidelity scaling. The portion of the scalable media objects is selectively delivered to adapt the QoS requirement. This framework can satisfy the various QoS requirements on bandwidth, even though the clients share a multipoint channel. By analysis on the rate-distortion characteristics of scalable media object framework, we shows our approach is feasible to support the various QoS requirements.

1. 서 론

통신망을 통한 멀티미디어 서비스가 확대됨에 따라 통신망을 통해 전달되는 데이터의 양이 폭발적으로 증가하게 되고, 이에 따라 전송 데이터의 증가율은 통신망 대역폭의 확대율을 항상 초과하게 된다. 이 영향으

로 통신망의 사용자는 자신이 기대하는 서비스 품질(QoS, Quality of Service)을 항상 보장받을 수는 없게 된다. 이러한 문제에 대응하기 위하여 통신망의 전송 계층에서 대역폭의 할당 관점에서 제한된 사용자의 QoS를 만족시키기 위한 연구가 많이 진행되어 왔으며, 허용제어(admission control)나 자원 예약(resource reservation)등에 관한 연구가 이 부류의 노력에 해당된다[1].

그러나 멀티미디어의 데이터 량이 매우 방대하여 전

[†] 정 회 원 : 한국전자통신연구원 컴퓨터·소프트웨어기술연구소 멀티미디어연구부 선임연구원

^{††} 정 회 원 : 한국SW진흥원 기술진흥부장

논문접수 : 1999년 3월 4일, 심사완료 : 1999년 10월 4일

송 계층의 대역폭 할당에 관한 노력만으로 사용자의 QoS를 만족시키기에는 한계가 있다. 특히 비디오의 경우에 문제가 심각해지며 MPEG 등의 비디오 압축기법이 적용되어도 여전히 방대한 데이터 량의 문제는 남아있다. 이에 따라 전송될 멀티미디어 데이터량을 조정함으로써 주어진 대역폭을 최대한 활용할 수 있도록 하는 미디어 스케일링(media scaling)에 관한 연구가 최근에 진행되고 있다[2, 3]. 미디어 스케일링은 다양한 QoS 요구에 따라 서로 다른 양의 데이터를 전송할 수 있도록 멀티미디어 데이터를 처리하는 기법으로서, 주어진 비디오 스트림에 대해서 주로 프레임 율의 조정이나 화질의 조정을 기반으로 한다.

멀티포인트 채널을 공유하는 여러 사용자의 QoS 요구가 서로 다른 경우에는 협상과정을 거쳐서 각 사용자가 공유할 특정 QoS 값이 결정되며, 이 협상 값에 대응되는 비트율을 가지는 멀티미디어 콘텐츠의 전송이 이루어진다. 이때, 협상 QoS 값은 모든 QoS 요구에 대한 타협점이기 때문에 실제 더 좋은 대역폭의 링크로 연결된 사용자는 낮은 화질 등의 서비스 제공에 대하여 만족하지 못하게 된다. 한편 더 낮은 대역폭의 링크로 연결된 사용자는 상대적으로 높아진 비트율의 서비스로 인한 추가 전송지연 등의 불만족을 겪게 된다. 이런 QoS 불만족 문제는 64 Kbps의 ISDN 망부터 155 Mbps의 ATM 망까지 다양한 대역폭의 링크가 연동되어 있는 이질적인 통신망 환경(heterogeneous networking environment)에서 더욱 심각해진다. 이러한 문제에 대해서 미디어 스케일링은 동일한 콘텐츠를 폭넓은 비트율의 스트림으로 제공할 수 있기 때문에 다양한 사용자를 모두 만족시킬 수 있는 가능성을 내포하고 있다. 그러나, Zeng[3]이나 Chan[4]의 방식처럼 특정 QoS 요구에 대해서 단순히 비트율을 조정하는 형태의 직관적인 미디어 스케일링 방식으로는 이질적인 통신망 환경의 멀티포인트 채널에서 발생하는 복잡한 QoS 요구 형태를 만족시킬 수 없다.

본 논문에서는 대역폭의 규모가 다양한 통신망 환경에서 QoS 적용형 멀티미디어 서비스에 적합한 스케일러블 미디어 객체 프레임워크를 제안한다. 이 프레임워크에서는 주어진 MPEG-2 비디오 비트스트림은 세밀하게 분할되어 스케일러블 미디어 객체의 집합으로 변환되고, 그 집합에 해당되는 스케일러블 미디어 객체들 사이에는 종속성이 존재한다. 주어진 비디오 비트스트림은 픽처 타입에 따른 스케일링 방식인 시간적 스케

일링(temporal scaling)과 DCT 계수의 범위에 따른 스케일링 방식인 충실도 스케일링(fidelity scaling)의 두 가지 분할 기법에 의하여 스케일러블 미디어 객체로 변환된다. 특정 QoS 요구를 가진 서비스 요청이 발생한 경우에는 미디어 서버에 존재하는 스케일러블 미디어 객체의 전체 집합 혹은 부분 집합이 클라이언트에게 전송된다. 그리고 이들은 클라이언트에서 재결합되어 하나의 MPEG-2 비디오에 호환되는 비트스트림으로 구성되어 MPEG-2 디코더에서 재생된다.

스케일러블 미디어 객체 프레임워크의 특성은 다음과 같다. 첫째, 주어진 MPEG-2 비디오의 비트스트림으로부터 매우 폭넓은 비트율의 비트스트림을 조합할 수 있는 스케일러블 미디어 객체를 제공함으로써 다양한 QoS 요구를 유연하게 적용시킬 수 있다. 둘째, 요구된 QoS에 적합한 비트율의 데이터만 전송함으로써 대역폭이 불필요하게 낭비되는 현상을 방지할 수 있다. 셋째, 서로 다른 대역폭을 가지는 링크들이 하나의 멀티포인트 채널을 공유하는 경우에도 각 세부 채널에서 서로 다른 스케일러블 미디어 객체를 전송함으로써 기존의 직관적인 미디어 스케일링의 문제를 해결할 수 있다. 넷째, 세션이 운영되는 동안 계속해서 QoS가 변화되는 일반적인 채널의 모델의 경우에도 동적인 QoS 적응을 지원할 수 있도록 쉽게 확장될 수 있다.

본 논문의 구성은 다음과 같다. 2절에서는 MPEG-2 비디오 코딩의 원리에 대하여 알아보고, 3절에서는 이질적 통신망 환경에서 직관적인 미디어 스케일링의 문제점으로 지적한다. 4절에서는 스케일러블 미디어 객체 프레임워크를 제안하고, 5절에서는 스케일러블 미디어 객체의 데이터 양에 따른 신호왜곡 특성을 분석한 후, 6절에서 결론을 맺는다.

2. MPEG-2 비디오의 개요

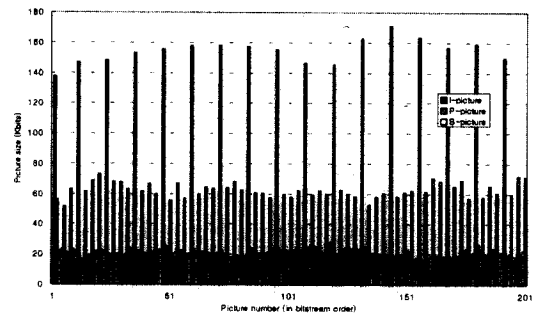
MPEG-2 비디오의 코딩 원리는 이동 보상 코딩과 블럭 기반 변환 코딩을 근간으로 한다[5]. MPEG-2 비디오 비트스트림은 단단계의 문법 구조로 표현된다. 최상위 단계의 문법 구조는 시퀀스(sequence)이다. 시퀀스는 일련의 코딩된 픽처(picture)로 구성되며, 이들 픽처들은 선택적으로 픽처 그룹(Group of Pictures, GOP)을 형성할 수 있다. 각 픽처는 슬라이스(slice), 매크로 블럭(macro block), 그리고 블럭(block)의 단계로 세분화된다. 블럭은 8x8 픽셀 배열이며 변환 코

덩의 단위로 이용된다. 매크로 블럭은 이동 보상 코딩의 단위로 이용되며, 2x2개의 밝기정보(luminance) 블럭과 이에 상응하는 색상정보(chrominance) 블럭의 집합으로 구성된다. 하나의 매크로 블럭에 포함되는 색상정보 블럭의 개수는 해당 비디오 시퀀스를 코딩할 때 적용된 색상형식(chroma format)에 따라서 달라진다. 한 픽처내에서 수평방향으로 연속된 매크로 블럭들은 하나의 슬라이스를 구성하며, 슬라이스는 데이터 손실의 파장을 줄이기 위한 재동기 단위로 사용된다.

각 8x8 픽셀 블럭의 픽셀 도메인 값들은 DCT(Discrete Cosine Transform)를 통하여 주파수 도메인의 값으로 변환된다. 주파수 도메인에서는 블럭의 대부분의 에너지가 소수의 저주파 항으로 집결되는 특성을 보이고 있으며, 이러한 저주파 항들은 블럭의 좌측상단에 위치하는 것들이다. 변환을 통한 DCT 계수들은 양자화되며, 이 과정에서 사람의 눈에 덜 민감한 영역인 고주파 항(블럭의 우측하단 부분)의 상당수의 값들은 0으로 된다. 양자화 과정(quantization)은 MPEG-2 비디오 코딩에서 유일하게 데이터 손실이 발생하는 부분이다. 2차원의 양자화된 DCT 계수들은 일정한 스캔 규칙에 따라 1차원 스트링으로 재배열되며, 이는 다시 Run-Length 코딩(RLC)을 거치게 된다. RLC는 그 값이 0이 아닌 DCT 계수의 값과 이 계수에 선행하는 0 값의 계수의 개수를 (런, 레벨)의 형태로 표현하는 코딩 방식이다. 비디오 코딩에서 이러한 방식을 채택한 동기는 일반적인 영상에 대하여 각 블럭을 양자화하고 나면 우측하단 영역의 고주파 항의 값들이 대부분 0이 된다는 특성 분석에 기인한다. 마지막으로 RLC된 코드는 통계적인 중복성을 줄이기 위하여 허프만 코딩 등의 가변길이 코딩이 적용된다.

MPEG-2 비디오 코딩은 시간축의 중복성을 제거하기 위하여 이동보상 코딩을 수행하며, 이의 적용방식에 따라 각 픽처는 I-픽처, P-픽처 혹은 B-픽처의 타입으로 코딩된다. I-픽처는 인트라 픽처라고도 하며 시간 축에서 다른 픽처를 전혀 참조하지 않고 독자적으로 코딩된 픽처이다. 즉, 이동보상 코딩이 적용되지 않고 공간적인 코딩만 적용된 픽처이며, 디코딩 될 때에도 독자적으로 디코딩 될 수 있는 픽처이다. P-픽처는 단방향 이동보상 코딩이 적용된 픽처이며, 이전의 I-픽처 혹은 P-픽처를 참조하여 이동보상이 수행된다. B-픽처는 과거와 미래의 I-픽처 혹은 P-픽처를 참조하여 양방향 이동보상 코딩이 적용된 픽처이다. 즉, I-픽처

는 독자적으로 디코딩이 될 수 있는 반면, P-픽처나 B-픽처는 다른 참조 픽처가 존재해야만 디코딩이 될 수 있다는 것을 의미한다. 다른 관점에서 보면, I-픽처는 원래 픽셀 값이 블럭 중심으로 코딩된 형태라고 볼 수 있으며, P-픽처나 B-픽처는 참조 픽처와의 차이값이 블럭 중심으로 코딩된 형태라고 볼 수 있다. 일반적으로 코딩된 I-픽처의 크기는 P-픽처보다 크고, P-픽처의 크기는 다시 B-픽처의 크기보다 큰 관계가 성립된다((그림 1) 참조).



(그림 1) I 타입, P 타입 및 B 타입 픽처들의 데이터 양 분포: 1.43 Mbps로 MPEG-2 코딩된 Flower Garden 비디오 시퀀스

3. 이질적 통신망 환경

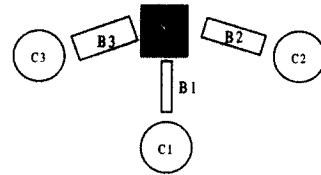
멀티미디어 서비스를 위하여 사용되고 있는 통신망의 환경을 살펴보면 매우 다양한 요소가 포함되어 있다. 통신망을 구성하는 이질적인 요소의 대표적인 것으로는 대역폭의 다양성과 단말기의 다양성을 들 수 있다. 첫째, 통신 링크 대역폭의 측면에서 보면 64 Kbps의 대역폭을 가지는 ISDN 선로부터 155 Mbps의 대역폭을 가지는 ATM 망까지 다양한 대역폭의 링크가 혼재되어 있다. 둘째, 단말기의 처리 능력을 살펴보면 디스플레이의 해상도의 차이를 비롯하여 간단한 규모의 소프트웨어 비디오 디코더만 사용가능한 단말기부터 전용 하드웨어 비디오 디코더가 장착된 고성능 단말기까지 다양한 처리 능력의 단말기가 함께 이용되고 있다. 이와 같은 다양한 통신망의 구성요소는 통신망 환경을 매우 이질적(heterogeneous)으로 만들고 있으며[6], 이로 인한 멀티미디어 서비스 사용자의 폭넓은 QoS 요구를 만족시키기 위한 노력이 요구된다.

이질적인 통신망 환경에서 사용자들의 다양한 QoS를 만족시키기 위한 연구가 활발히 진행되어 왔으며,

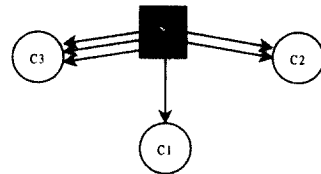
다음과 같이 두 부류로 구분할 수 있다. 한 그룹은 QoS 요구사항을 채널의 주어진 용량에 맞추기 위하여 대역폭 할당이나 허용 제어, 자원 예약 등의 기법을 연구하고 있다[4, 7]. 또 다른 그룹은 미디어 스케일링 기법을 통하여 멀티미디어 데이터의 데이터량을 조절함으로써 채널에 실리는 부하를 경감시키려는 노력을 하고 있다[2, 3, 8]. 이 두 가지 접근 방식은 각각 트랜스포트 시스템 계층과 응용 계층에서 QoS를 만족하기 위한 노력으로서 의미를 가지고 있으며, 보다 진보된 멀티미디어 통신 서비스를 제공하기 위하여 이 두 가지 접근 방식을 혼합하여 적용할 수도 있다. 본 논문은 이 중에서 후자의 접근방식인 미디어 스케일링과 관련된 부분을 다루고 있다.

미디어 스케일링은 주로 비디오 데이터에 대해서 적용되고 있으며, 이는 비디오 데이터의 경우에는 일부 데이터가 손실되어도 복구한 후의 결과가 사람에게 덜 민감한 특성을 가지기 때문이다. 즉, 미디어 스케일링은 비디오 코딩의 원리를 이용하여 비디오 데이터의 일부분을 제거함으로써 주어진 QoS 요구사항에 적합하도록 발생 데이터율을 조정하는 것이다. MPEG-2 비디오의 미디어 스케일링에 이용되는 대표적인 방법으로는 프레임율 조정(frame rate control)과 충실도 조정(fidelity control)을 들 수 있다. 프레임율 조정은 수신 측에서 비트스트림의 디코딩 시의 중요도가 낮은 프레임을 생략하고 전송하는 방식이다. 충실도 조정은 DCT 계수 중에서 중요도가 낮은 계수인 AC항의 일부를 생략하고 전송하는 방식이다. 영상의 처럼 데이터가 실시간으로 발생하는 서비스에서는 비디오 인코딩의 입력 변수인 양자화 계수를 변화시키거나 발생 프레임율을 인코더에서 제어함으로써 원하는 QoS에 대한 데이터율을 인코더로부터 얻을 수 있다. 그러나, VOD(Video On Demand)나 하이퍼미디어 등과 같은 저장형 서비스에서는 이미 코딩되어 저장된 MPEG-2 비트스트림을 다루게 되므로, 데이터율의 조절을 위하여 인코딩시의 매개변수를 조절하는 방식은 불가능하다. 따라서, 저장형 서비스를 위해서는 별도의 미디어 스케일링 기법이 필요하다.

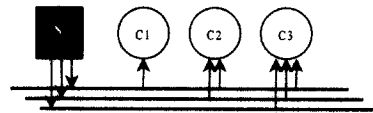
코딩되어 있는 MPEG-2 비트스트림을 스케일링하기 위한 직관적인 방식은 서로 비트율이 다른 여러 버전의 비트스트림을 미디어 서버에 유지하는 방법이다. 이 방식을 따르면 동일한 내용의 비디오가 상당부분 각 버전에 걸쳐 중복되어 있게 되며, 이는 곧 미디어



(a) 각 클라이언트 C_i 가 서로 다른 대역폭 B_i 로 미디어 서버 S 와 연결된 상태 (이때, $B_3 > B_2 > B_1$)



(b) 서로 다른 QoS 요구를 가지는 세 개의 점대점 채널 설정



(c) 서로 다른 QoS 요구를 가지고 하나의 멀티포인트 채널을 공유하는 경우

(그림 2) 다양한 대역폭의 이질적인 통신망 환경

서버의 불필요한 저장공간 낭비를 초래한다. 직관적인 미디어 스케일링 기법은 각 버전의 비트스트림이 모두 독립적으로 디코딩 될 수 있는 형태(self-contained 형태)로 구성되어 있기 때문에 각 버전마다 중복되는 내용의 데이터가 포함되는 것이다. 또한 이 방식은 이질적인 통신망 환경에서 멀티포인트 멀티미디어 서비스를 제공하는데 상당한 제약을 초래한다. (그림 2)는 이와 같은 직관적인 미디어 스케일링 방식의 한계를 설명하고 있다.

(그림 2)-(a)는 미디어 서버 S 와 클라이언트 C_1 , C_2 및 C_3 가 이질적인 통신망 환경에서 연결되어 있는 상황을 보이고 있다. 이때, 서버와 클라이언트 C_1 , C_2 및 C_3 를 연결하는 각 링크의 대역폭을 각각 B_1 , B_2 , 그리고 B_3 라 했을 때, 각 대역폭의 관계는 $B_3 > B_2 > B_1$ 라고 가정한다. 즉, C_3 는 가장 넓은 대역폭의 링크로 연결되어 있고, C_1 은 가장 좁은 대역폭의 링크로 연결된 경우이다. (그림 2)-(b)는 각 클라이언트가 미디어 서버와 점대점 세션을 독립적으로 형성하고, 각자의 대역폭에 해당되는 QoS에 따라 MPEG-2 비디오 클립을 요구하는 상황을 보이고 있다. 이 경우에는 각 세션의

대역폭이 다르지만 각 세션이 독립적으로 운영되고 있기 때문에 직관적인 미디어 스케일링 방식으로도 서비스가 가능하다. 즉, 각 세션마다 그의 대역폭을 만족시킬 수 있는 비트율 버전의 비트스트림을 전송함으로써 모든 클라이언트의 QoS 요구사항을 만족시킬 수 있다. 물론 이때에도 미디어 서버의 저장공간의 낭비 문제는 해결되지 않은 채 남아있다.

그러나 (그림 2)-(c)의 경우처럼 모든 클라이언트가 동일한 멀티캐스트 세션에 참여하고 있는 경우에는 여러 버전의 비트스트림을 유지하는 직관적인 방식은 더욱 한계를 가지게된다. 어느 특정 링크의 대역폭(예를 들어 C2의 링크)에 적합한 버전의 비트스트림만을 전송하는 경우에는 C3처럼 충분한 대역폭의 링크를 가지고도 저 수준의 서비스를 받거나 혹은 C1의 경우처럼 불필요하게 대역폭을 낭비하게되는 링크가 발생한다. 본 논문에서는 이러한 직관적인 미디어 스케일링 기법의 문제점을 극복할 수 있는 스케일러블 미디어 객체 프레임워크를 제안한다.

4. 스케일러블 미디어 객체 프레임워크

스케일러블 미디어 객체 프레임워크(scalable media object framework)은 분할된 MPEG-2 미디어 객체를 표현하는 데이터 모델이다. 하나의 MPEG-2 비디오 비트스트림은 세부 비트스트림의 집합으로 분할되며, 이러한 세부 비트스트림을 스케일러블 미디어 객체라 부른다. 스케일러블 미디어 객체로 분할하는 기법으로서 시간적 스케일링과 충실도 스케일링이 적용되고 있다.

4.1 시간적 스케일링

일반적으로 시간적 스케일링(temporal scaling)은 주어진 시간에 전송되는 비디오 프레임의 수를 줄임으로써 비디오 스트림의 시간적 해상도를 줄이는 방식을 의미한다. 그러나, MPEG-2 비디오 비트스트림에 포함되어 있는 각 픽처가 모두 독립적 형태인 인트라 코딩된 것이 아니기 때문에 임의의 프레임울 삭제하는 형태의 단순한 프레임 울 조정 방식은 적용이 불가능하다. 따라서, 본 논문에서는 픽처 타입에 따른 프레임 울 조정 방식을 사용한다.

시간적 스케일링은 MPEG-2 비디오 비트스트림으로부터 픽처 계층을 픽처 타입에 따라 분할하고, 이들을 다시 MPEG-2 비디오 비트스트림으로 복원하는 과정

을 의미한다. 시간적 스케일링 분할과정을 통하여 하나의 MPEG-2 비디오 비트스트림은 I-서브스트림, P-서브스트림 및 B-서브스트림의 세 개의 중간 표현구조로 나누어진다. 이 서브스트림들은 각기 한가지 타입의 픽처들의 집합이다. 즉, I-서브스트림은 I-픽처들을 모아놓은 것이고, P-서브스트림은 P-픽처들을 모아놓은 것이다. MPEG-2 비디오의 계층적 문법 구조 중에서 픽처 계층으로부터 블럭 계층까지의 부가정보인 헤더 등의 정보는 각 I-, P- 및 B-서브스트림에 표현되어 있다. 한편 픽처 계층보다 상위 계층의 부가정보인 시퀀스 헤더와 GOP 헤더 정보는 I-서브스트림에만 포함되어 있다. 따라서 I-서브스트림은 MPEG-2 비디오의 신택스와의 호환성을 가지며, 독자적으로 디코딩 될 수 있다. P-서브스트림과 B-서브스트림은 상위 계층 헤더 정보를 가지고 있지 않고, 이동보상 코딩의 영향으로 독자적으로 디코딩 될 수 없는 불완전한 형태를 지닌다. P-서브스트림이 디코딩되기 위해서는 I-서브스트림과 결합이 되어야하며, 이 결합은 각 픽처간의 비트스트림 순서(bitstream order)에 의거하여 이루어지게 된다. 이때, P-서브스트림은 I-서브스트림에 종속된다고 한다. B-서브스트림의 디코딩을 위해서는 I-서브스트림과 P-서브스트림이 모두 필요하며, B-서브스트림은 I-서브스트림과 P-서브스트림에 종속된다고 한다.

4.2 충실도 스케일링

충실도 스케일링은 MPEG-2 비디오 문법 구조의 블럭 계층을 코딩할 때에 산출된 DCT 계수의 개수를 조정함으로써 화질을 저하시키는 미디어 스케일링 기법이다. 일반적인 영상의 경우에 고주파항의 DCT 계수가 복원된 화질에 미치는 영향은 매우 미미한 특성을 보인다.

8x8 픽셀의 2차원 블럭을 변환한 64개의 2차원 DCT 계수는 zigzag 스캐닝이나 교체(alternate) 스캐닝 과정을 통하여 1차원 계수열로 변환된다. 64개의 1차원 DCT 계수는 N개의 계수 그룹 G_i 로 분할되며 이 계수 그룹들은 순서 있는 집합 (ordered set)인 S에 포함된다. 이때, N 값의 범위는 $0 < N \leq 64$ 이며, i 값의 범위는 $1 \leq i \leq N$ 이다. 순서 있는 집합 S의 멤버의 갯수인 N은 DCT 계수 집합 S가 표현할 수 있는 충실도의 계층 수를 나타낸다. S의 멤버인 계수 그룹 G_i 는 k개의 연속된 DCT 계수를 가지고 있으며, 임의의 특정

DCT 계수는 $i \neq j$ 인 서로 다른 계수 그룹 G_i 와 G_j 에 동시에 포함될 수 없다.

$i < j$ 인 경우에 G_i 는 G_j 보다 더 저주파향 성향의 계수를 가지고 있는 그룹이며, G_i 는 G_j 에 선행되었다고 한다. 이는 연속한 그룹간의 종속성을 나타내는 것이며, 그룹 G_i 는 그룹 G_j 에 종속된다고 한다. 그리고, 한 블록에서의 최저주파향인 DC 항은 항상 G_1 그룹에 속하며, G_1 그룹의 첫 번째 요소이다. 즉, 그룹 G_i 가 디코딩되기 위해서는 그룹 G_i 및 G_i 를 선행하는 모든 그룹이 연속적으로 비트스트림을 구성해야 된다.

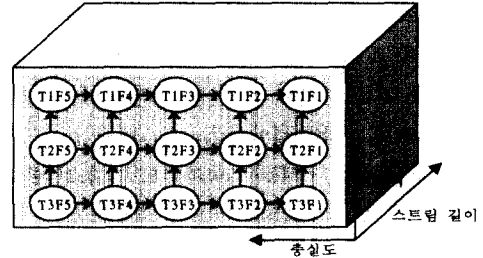
총실도 스케일링은 인트라 픽처뿐만 아니라 인터 픽처에도 적용할 수 있다. 이는 DCT가 인트라 픽처의 밝기 신호를 변환하는데 적용될 뿐 아니라, 인터 픽처의 이동추정 예측 오차를 변환하는데도 적용되기 때문이다. 따라서, 총실도 스케일링은 시간적 스케일링의 결과인 I-서브스트림, P-서브스트림 및 B-서브스트림에 모두 적용될 수 있다.

4.3 스케일러블 미디어 객체

주어진 MPEG-2 비디오 비트스트림에 시간적 스케일링과 총실도 스케일링이 적용되면 그 결과로서 미디어 객체의 집합이 생성되며, 이러한 미디어 객체를 스케일러블 미디어 객체라 한다. 각 스케일러블 미디어 객체가 포함하고 있는 비디오 데이터의 관점에서 보면 스케일러블 미디어 객체는 서로 내용 중복성으로부터 독립적이다. 즉, 어느 스케일러블 미디어 객체도 다른 스케일러블 미디어 객체와 중복되는 내용을 표현하고 있지 않다. 그 이유는 주어진 MPEG-2 비디오 비트스트림에서 대해서 상호 직교 분할 방식인 시간적 스케일링과 총실도 스케일링을 적용한 결과로서 한 집합의 스케일러블 미디어 객체가 얻어졌기 때문이다.

스케일러블 미디어 객체는 시간적 스케일링과 총실도 스케일링을 모두 적용한 결과이기 때문에 각 스케일링의 결과에서 나타나는 데이터 종속성은 복합적인 형태로 스케일러블 미디어 객체에 계승되어 적용된다. 이에 따라, 주어진 MPEG-2 비디오 비트스트림에 미디어 스케일링을 적용함으로써 해당 MPEG-2 비디오 비트스트림을 2차원 종속성 그래프로 표현할 수 있게 된다. 이때, 2차원 종속성 그래프의 노드는 각 스케일러블 미디어 객체이며, 노드간의 링크는 시간적 종속성과 총실도 종속성을 의미한다. 이러한 2차원 종속성

그래프는 주어진 MPEG-2 비디오 비트스트림에 대한 스케일러블 특성의 내포하는 데이터 모델이며, 이를 스케일러블 미디어 객체 프레임워크라고 부른다.



(그림 3) 스케일러블 미디어 객체 프레임워크

(그림 3)은 스케일러블 미디어 객체 프레임워크를 보이고 있다. 각 노드 $t_i f_j$ 는 시간적 스케일링의 i -서브스트림과 총실도 스케일링의 계수 그룹 G_j 에 해당되는 스케일러블 미디어 객체를 나타낸다. 이때, i 의 값은 I-서브스트림, P-서브스트림 및 B-서브스트림에 대해서 각각 1, 2 및 3의 값을 가진다. (그림 3)은 총실도 스케일링을 적용하여 5개의 계수 그룹만 형성한 경우 ($1 \leq j \leq N, N=5$)를 보이고 있다. 비록 G_j 는 임의의 개수의 연속된 1차원 DCT 계수를 가질 수 있지만, 본 논문에서는 이를 단순히 다음의 할당규칙을 적용하였다. 즉, $G_1=(DC)$, $G_2=(AC1, AC2)$, $G_3=(AC3, AC4, AC5)$, $G_4=(AC6, AC7, AC8, AC9)$, 그리고 $G_5=(AC10, AC11, \dots, AC63)$. 이 할당규칙은 다음절에서 데이터 양에 따른 신호의 왜곡 특성 (rate-distortion characteristics)를 분석하는 과정에서 동일하게 적용된다.

노드 $t_1 f_1$ 은 DC항의 계수만으로 하나의 블록이 표현된 I 타입의 픽처와 상위 계층 헤더로 구성된 비트스트림을 의미한다. 다시 말하면, $t_1 f_1$ 은 $t_1 f_F$ 으로 표현되는 I-서브스트림에 총실도 스케일링을 적용하여 DC항의 값만으로 각 블록을 대표하도록 하는(이를 $t_T f_1$ 으로 표현함) 비트스트림이다. 이때, I-서브스트림은 독자적 표현형태²⁾(self-contained)이고, DC항도 역시 독자적 표현형태이기 때문에 $t_1 f_1$ 가 독자적 표현형태임은 명확하다. 이 사실은 노드 $t_1 f_1$ 이 스케일러블 미디어 객체 프레임워크에서 다른 스케일러블 미디어 객체

2) 독자적 표현형태의 비트스트림은 추가적인 다른 정보 없이도 MPEG-2 비디오 디코더를 통하여 디코딩될 수 있다.

가 디코딩되는 과정에서 가장 필수적으로 필요한 기본 객체임을 의미한다.

노드 $t_1 f_2$ 는 I 타입의 픽처에 대해서 각 블록의 DCT 계수 중 {AC1, AC2}항만을 모아놓은 비트스트림이다. 노드 $t_2 f_1$ 는 P 타입의 픽처로 구성된 비트스트림인데, 각 픽처의 블록들은 DC항의 계수로만 표현되어 있다. 기본 객체인 $t_1 f_1$ 은 독자적으로 디코딩 가능한 독자적 표현형태이지만, $t_1 f_2$ 이나 $t_2 f_1$ 등의 스케일러블 미디어 객체들은 독자적으로 디코딩이 될 수 없는 불완전한 형태이다. 이러한 불완전한 스케일러블 미디어 객체들은 디코딩 될 수 있는 형태로 되기 위해서는 기본 객체 등과 같은 다른 스케일러블 미디어 객체들과의 결합이 필요하게되며, 이는 객체들간의 종속성을 의미한다.

(그림 3)에서 두 노드를 연결하는 화살표 링크는 객체들간의 종속성을 표현하고 있다. 화살표 링크의 원점 노드의 객체 (예를 들어, $t_1 f_2$)가 독자적 표현형태가 되어 디코딩될 수 있기 위해서는 링크의 목적지 노드 객체 (예를 들어, $t_1 f_1$)가 필요하다는 것을 의미한다. 시간적 스케일링 클래스 2와 충실도 스케일링 클래스 2를 만족하는 MPEG-2 비디오 비트스트림을 구성하기 위해서는 $t_2 f_2$, $t_2 f_1$, $t_1 f_2$ 및 $t_1 f_1$ 객체들이 모두 필요하며, 이들을 스케일러블 미디어 객체 그룹 T2F2라 부른다. 따라서, 특정 비트율을 가지는 MPEG-2 비디오 호환 비트스트림을 구성하기 위해서는 비트율을 만족시키는 스케일러블 미디어 객체 그룹 T i F j 가 필요하다. 그리고 T i F j 를 형성하기 위해서는 $i \cdot j$ 개의 스케일러블 미디어 객체 $t_m f_n$ ($m \leq i, n \leq j$)들이 필요하다.

4.4 이질적 통신망 환경에서 스트리밍 서비스에 대한 QoS 적응성

통신망을 통한 멀티미디어 서비스 환경에서 미디어 서버에는 MPEG-2 비디오 비트스트림들이 스케일러블 미디어 객체 프레임워크로 표현되어 있고, 클라이언트들의 다양한 QoS 요구에 따라 적절한 스케일러블 미디어 객체들을 선택하여 전송, 결합, 디코딩되는 과정을 거치게된다. 클라이언트가 요구한 QoS 값을 비트율로 환산하고, 환산된 비트율에 해당되는 스케일러블 미디어 객체 그룹을 선정한다. 특정 스케일러블 미디어 객체 그룹이 선정되면, 그에 포함되는 스케일러블

미디어 객체를 전송함으로써 클라이언트의 QoS 요구를 만족시킬 수 있게된다.

이에 따라, 클라이언트들이 이질적 통신망 환경에서 어떤 형태로 미디어 서버에 연결되어 있는지 무관하게 클라이언트로부터 발생하는 다양한 QoS 요구에 적용할 수 있는 기반을 스케일러블 미디어 객체 프레임워크가 제공한다. 앞 장의 (그림 2)-(b)에서 보인 바와 같이 각 점대점 세션들이 서로 다른 QoS를 요구하는 경우에는 서로 다른 스케일러블 미디어 객체 그룹을 각 세션에 전송함으로써 요구된 QoS를 만족시킬 수 있다. Zeng[3]이나 Chan[4]의 방식과 같은 직관적인 미디어 스케일링 방식에서도 각 QoS를 만족시킬 수는 있었으나, 중복된 데이터에 따른 저장장치의 낭비가 발생하게된다. 그러나, 스케일러블 미디어 객체 프레임워크를 사용하는 경우에는 데이터의 중복이 발생하지 않기 때문에 불필요한 저장장치의 낭비를 막을 수 있다.

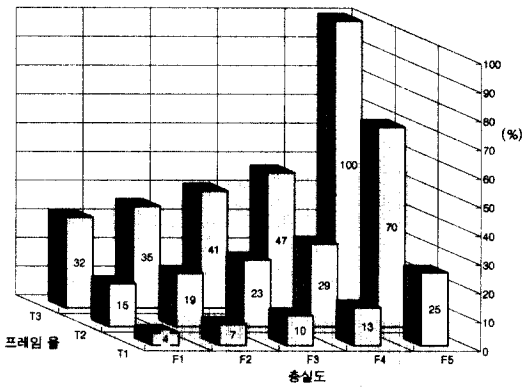
클라이언트들이 동일한 멀티캐스트 세션을 공유하는 상황인 (그림 2)-(c)에서는 Zeng[3]이나 Chan[4]등과 같은 직관적인 미디어 스케일링 방식으로는 각 QoS를 만족시킬 수 없는 경우를 보이고 있다. 이러한 경우에 본 논문에서 제안한 스케일러블 미디어 객체 프레임워크를 적용하면 각 클라이언트의 QoS를 모두 만족시킬 수 있다. 예를 들어, 클라이언트 C1에게는 스케일러블 미디어 객체 그룹 T1F1을 전송하고, C2에게는 T2F2를 전송하며 C3에게는 T3F3를 전송할 수 있다. 이 경우에 C1에게는 t1f1 객체만 전송되고, C2에게는 t1f1, t2f1, t1f2 및 t2f2가 전송되는데 t1f1은 두 클라이언트에게 동일한 멀티캐스트 세션을 통해서 전달되므로 불필요한 대역폭을 낭비하지 않게 된다.

5. 데이터 양에 따른 신호 왜곡 특성

본 절에서는 스케일러블 미디어 객체에 대하여 데이터 양에 따른 신호 왜곡 특성(rate-distortion characteristics)를 분석한다. 시험용 MPEG-2 비디오 시퀀스는 Flower Garden이며 1.43 Mbps로 인코딩되어 있다. 이 시퀀스의 픽처 패턴은 2절의 (그림 1)에 보인 바와 같다. 시험용 MPEG-2 비디오 시퀀스는 선처리 분할과정을 통하여 t1f1에서 t3f5까지의 15개의 스케일러블 미디어 객체로 변환되었다.

(그림 4)는 각 스케일러블 미디어 객체 그룹의 전송에 필요한 데이터 양을 미디어 스케일링 이전의 원래

MPEG-2 비디오 시퀀스의 데이터 양에 대한 백분율로 나타낸 그래프이다. 요구된 QoS에 대하여 스케일러블 미디어 객체 그룹 T1F1이 선택된 경우라면 원래 비트 스트림의 4%만 전송된다. T2F4 그룹이 선택된 경우에는 원래 비트스트림의 61%만 전송되며, 이는 스케일러블 미디어 객체 $t_i f_j (i \leq 2, j \leq 4)$ 들의 데이터 양을 합산한 결과이다. (그림 4)를 통하여 스케일러블 미디어 객체 프레임워크가 데이터 양을 얼마나 다양하게 조절할 수 있으며, 결과적으로 다양한 QoS 요구에 적절히 적용시킬 수 있게된다.



(그림 4) 압축된 Flower Garden 시퀀스에 대한 각 스케일러블 미디어 객체 그룹의 데이터 양의 백분율

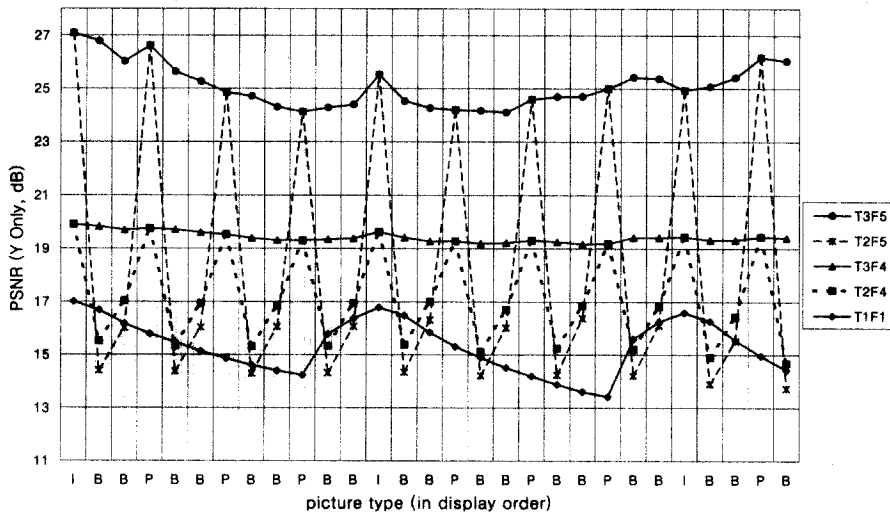
시간적 스케일링에 따른 픽처의 생략과 각 픽처의 충실도 스케일링에 따른 신호 왜곡 특성을 신호대 잡음비 (PSNR, Peak Signal to Noise Ratio)로 비교한 결과가 (그림 5)에 나타나 있다. PSNR은 일부 스케일러블 미디어 객체 그룹을 이용하여 디코딩한 스트림의 충실도를 압축 이전의 원본 스트림과 비교한 것이다. 한 픽처의 PSNR은 식 (1)을 통하여 얻어진다.

$$PSNR = 10 \log \left(\frac{\delta^2}{MSE} \right), \quad (1)$$

$$MSE = \frac{1}{M \times N} \sum_{m=1}^M \sum_{n=1}^N |L(m, n) - L'(m, n)|^2$$

이때, 영상의 크기는 $M \times N$ 픽셀, $L(m, n)$ 은 디코딩된 영상의 밝기 신호, $L'(m, n)$ 은 압축 이전의 원본 영상의 밝기 신호이며 δ 는 255이다.

(그림 5)에서 T3F5에 대한 그래프는 MPEG-2로 압축된 비디오 비트스트림이 압축되기 이전의 원본 신호에 대한 PSNR을 보이고 있으며, 이는 미디어 스케일링의 적용여부와 무관하게 압축 비트스트림을 디코딩한 후에 얻을 수 있는 충실도의 최대 값을 의미한다. T2F5에 대한 그래프는 I-픽처와 P-픽처의 위치에서 T3F5의 그래프와 교차한다. 이는 T2F5에서 B-픽처가 삭제된 것을 제외하면 T3F5와 동일한 것이기 때문이다. 이 삭제된 B-픽처들은 클라이언트 측에서 비트스



(그림 5) 압축 이전의 Flower Garden 시퀀스에 대한 스케일러블 미디어 객체 그룹의 충실도 비교

트림 복원 시에 이전 픽처와 동일한 신호를 가지도록 조정되었다. T2F5는 B-픽처 영역에서 T3F4보다 충실도가 저하된다. 이는 시간적 스케일링의 결과로 발생한 픽처 삭제의 영향이 충실도 스케일링의 결과인 DCT 계수의 삭제의 영향보다 신호 왜곡이 크다는 것을 의미한다.

화질의 주관적인 판단을 위하여 몇 가지 충실도 그룹에 따라 복원된 비트스트림에서 동일한 위치의 I-픽처³⁾를 선택하여 (그림 6)에 보였다. (그림 6)에서 표현된 화질들은 본 논문에서 제안한 스케일러블 미디어 객체 프레임워크를 통해서만 얻어질 수 있는 화질의 분포가 아니며, 이들은 단순히 각 픽처의 DCT 계수의 일부만을 택하여 복원할 수 있는 일반적인 화질이다. 단, 주관적인 비교에 필요한 화질 분포를 데이터 양의 절감 정도와 신호 왜곡도를 동시에 비교검토할 목적으로 제시되었다. (그림 6)-(a)에 보이는 TXF5는 모든 DCT 계수를 포함하고 있으며, 이 비트스트림은 미디어 스케일링에 사용한 MPEG-2로 압축된 비트스트림과 동일하다. 따라서, 이 영상의 화질은 미디어 스케일링을 적용하지 않은 비트스트림의 화질과 동일한 것이다 (PSNR값은 26dB). (그림 6)-(b)의 TXF1은 DC항만을 포함하고 있는데 데이터 양은 4%~32% 정도로 줄일 수 있는 반면((그림 4)의 그래프 참조), PSNR값은 16.8dB로 매우 낮아서 화질은 육안으로도 매우 열악한 특성을 보이고 있다. (그림 6)-(c)의 TXF3은 데이터 양을 10%~41%로 줄일 수 있는 반면, PSNR값은 18.6dB 정도이다. TXF4는 데이터 양을 13%~47%로 줄이면서 PSNR값은 19.6dB 정도이다. 그런데 (그림 6)-(d)에서 보이는 바와 같이 육안으로는 화질의 저하를 거의 인식할 수 없는 정도의 충실도를 제공하고 있다.

데이터 양에 따른 신호왜곡 특성은 실험에 사용된 비디오 시퀀스에 따라 다를 수 있기 때문에, 본 절에서 분석한 결과의 수치가 그 자체로서 절대적인 의미를 갖지는 못한다. 그러나 이러한 분석 수치의 상대적인 특성은 스케일러블 미디어 객체 프레임워크 이질적인 멀티미디어 통신망 환경에서 다양한 QoS에 대한 적응성을 풍부하게 제공할 수 있음을 명확히 보이고 있다. 즉, 시간적 스케일링과 충실도 스케일링을 복합적으로 적용하여 하나의 주어진 MPEG-2 비디오 비트스트림에 대하여 매우 폭넓은 범위의 비트 율을 가지는 데이터 집합을 생성할 수 있게 된다.

(그림 6) 다양한 충실도 스케일링에 따라 복원된 동일한 I-타입 픽처의 주관적 비교

3) (그림 5)의 그래프에서 두 번째 I-픽처를 비교한 것임

6. 결 론

본 논문에서는 다양한 QoS 요구에 적용할 수 있는 잘 정의된 미디어 스케일링 프레임워크의 필요성을 지적하였다. 이에 따라, 본 논문에서는 이를 만족시킬 수 있는 스케일러블 미디어 객체 프레임워크를 제안하였다. 또한 MPEG-2 비디오 비트스트림로부터 스케일러블 미디어 객체로 변환할 수 있는 변환기와 역변환기를 설계하고 구현하였다. 또한 시험용 비디오에 대하여 데이터 양에 대한 신호왜곡 특성을 분석함으로써 제안한 방식의 가능성을 확인하였다. 현재 진행하고 있는 연구 내용으로는 요구된 QoS 값으로부터 적절한 스케일러블 미디어 객체의 부분집합을 선택하는 알고리즘을 설계하는 부분과 동적으로 QoS 값이 변하는 경우에 최적으로 적용할 수 있도록 프레임워크를 확장하는 부분이 포함되어 있다.

참 고 문 헌

[1] P. Ferguson, G. Huston, *Quality of Service : Delivering QoS on the Internet and in Corporate Networks*, Wiley Computer Publishing, 1998.

[2] L. Delgrossi, C. Halstrick, D. Hehmann, et. al., "Media scaling in a multimedia communication system," *ACM Multimedia Systems*, Vol.2, pp.172-180, 1994.

[3] W. Zeng, B. Liu, "Rate shaping by block dropping for transmission of MPEG-precoded video over channels of dynamic bandwidth," *ACM Multimedia96*, Boston MA USA, pp.385-393, 1996.

[4] D.L. S. Chan, S.T. Chanson, "Scalability support for multiparty multimedia communications," *ACM Multimedia Systems*, Vol.6, pp.75-87, 1998.

[5] ISO/IEC JTC1/SC29/WG11, *Coding of moving pictures and associated audio -- Video part*, ISO/IEC 13818-2, Draft International Standard, March

1994.

[6] A. Banerjea, "Heterogeneous networking," *IEEE Multimedia*, Vol.4, No.2, pp.84-87, 1997.

[7] M. Krunch, S.K. Tripathi, "Impact of video scheduling on bandwidth allocation for multiplexed MPEG streams," *ACM Multimedia Systems*, Vol.5, pp.347-357, 1997.

[8] A.T. Campbell, G. Coulson, D. Hutchison, "Transporting QoS adaptive flows," *ACM Multimedia Systems*, Vol.6, pp.167-178, 1998.



김 형 철

e-mail : hckim@etri.re.kr

1986년 인하대학교 전자계산학과 (학사)

1988년 한국과학기술원 전산학과 (공학석사)

1995년 한국과학기술원 전산학과 (공학박사)

1988년~현재 한국전자통신연구원 컴퓨터·소프트웨어 기술연구소 멀티미디어연구부 선임연구원
관심분야 : Video Media Processing, Multimedia Group Communication, Parallel Processing of Video Coding



정 찬 근

e-mail : chan@software.or.kr

1979년 한국항공대학 전자공학과 (학사)

1981년 서울대학교 대학원 전자공학과(공학석사)

1983년~1999년 한국전자통신연구원 근무

1999년~현재 한국SW진흥원 기술진흥부장
관심분야 : Multimedia Groupware, Image Processing