

QoS 적응형 MPEG 비디오의 미디어 스케일링 방안 고찰

김 형 철*

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 영향으로 통신망의 사용자는 자신이 기대하는 서비스 품질 (QoS, Quality of Service) 을 항상 보장받을 수는 없게된다. 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[Ferg98, Beck98].

그러나 멀티미디어의 데이터 량이 매우 방대하여 전송 계층의 대역폭 할당에 관한 노력만으로 사용자의 QoS를 만족시키기에는 한계가 있다. 특히 비디오의 경우에 문제가 심각해지며 비디오 코딩이 적용되어도 여전히 방대한 데이터 량의 문제는 남아있다. 이러한 문제를 해결하기 위하여 다음의 두 가지 접근방법이 시도되고 있다.

첫째, 데이터 양이 많은 비디오의 경우에는 다양한 통신망 및 단말 환경 등의 요구를 만족시키기 위하여 계층적 비디오 코딩 (layered video

coding)에 관한 기술 개발이 활발히 진행되어 왔다 [Belz94, Han97]. MPEG-1 [Mpeg93]이나 H.261 [H261]등과 같은 기존의 비디오 코딩에서는 비디오 데이터 전체를 단 하나의 비트스트림으로 변환하는 방식을 기반으로 하고 있다. 그러나, 계층적 비디오 코딩은 비디오 데이터를 하나의 기본 계층 (basement layer)과 여러 개의 상위 계층 (enhancement layer)으로 변환하는 방식을 기반으로 하고 있다. 이때, 복원하는 단말측의 환경에 따라 기본 계층만 복원하여 최소 수준의 비디오만을 재생하던지, 혹은 기본 계층과 상위 계층을 함께 복원하여 더 나은 수준의 비디오를 재생하는 방식을 채택할 수 있다. MPEG-2 [Mpeg94]는 계층적 비디오 코딩 개념을 도입하여 SNR 스케일러빌리티, 공간 스케일러빌리티, 시간 스케일러빌리티 및 데이터 분할 등의 4가지 확장성을 지원하고 있다. 이때, 하나의 기본 계층과 4종류의 상위 계층으로 구성되며, 상위 계층은 응용에 따라 취사 선택 혹은 혼합하여 적용할 수 있도록 되어있다. 그러나, 비디오의 복원을 위하여 필수적으로 필요한 기본 계층의 데이터 량도 이미 방대하여 대역폭이 작은 이동 단말이나 전화선 등이 포함된 다양한 통신망 환경에서는 서비스 품질의 저하를 피할 수 없으며 이는 응용의 폭을 제한하게 된다. 계층적 비디오 코딩과 실시간 멀티미디어 통신을 함께 고려하여, 압축된 비디오 정보의 각 계층을 서로 다른 신뢰도를 가지는 채널로 전

*정회원, 한국전자통신연구원, 컴퓨터·소프트웨어기술연구소 멀티미디어연구부

송함으로써 전송 손실이나 지연으로 인한 서비스 질의 저하를 완화시키는 것이 가능하다. 그러나, 복원을 위하여 필수적으로 필요한 기본 계층의 데이터 량도 이미 방대하여 대역폭이 작은 이동 단말이나 전화선 등이 포함된 다양한 통신망 환경에서의 서비스 품질의 저하에 관한 문제는 여전히 해결되지 않고 있다.

둘째, 전송될 멀티미디어 데이터 량을 조정함으로써 주어진 대역폭을 최대한 활용할 수 있도록 하는 미디어 스케일링(media scaling)에 관한 연구가 최근에 진행되고 있다[Delg94, Zeng96]. 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. 넓은 의미로 해석하면 계층적 비디오 코딩에서 기본 계층과 상위 계층을 나누는 방법도 미디어 스케일링의 범주에 해당된다. 그러나, 앞에서 지적한 바와 같이 계층적 비디오 코딩이라 하더라도 기본 계층의 데이터 량이 상당히 많기 때문에, 이 기본 계층의 데이터를 세분화하는 과정이 필요하며, 최근의 미디어 스케일링에 관한 연구는 이 부분에 해당된다.

한편, 멀티포인트 채널을 공유하는 여러 사용자의 QoS 요구가 서로 다른 경우에는 협상과정을 거쳐서 각 사용자들이 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 불만

족을 겪게된다. 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[Zeng96]이나 Chan[Chan98]의 방식처럼 특정 QoS 요구만 만족시킬 수 있는 단순한 비트율 조정 형태의 직관적인 미디어 스케일링 방식으로는 이질적인 통신망 환경의 멀티포인트 채널에서 발생하는 복잡한 QoS 요구 형태를 만족시킬 수 없다.

따라서 보다 체계적인 미디어 스케일링 방식이 필요하며, 그 결과로 생성되는 미디어 데이터 조각들로부터 사용자가 요구하는 QoS의 미디어 비트스트림을 복원할 수 있는 표현 모델이 필요하다.

2. 멀티미디어 통신 서비스 모델

멀티미디어 서비스를 위하여 사용되고 있는 통신망의 환경을 살펴보면 매우 다양한 요소가 복합되어있다. 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 단말

가까지 다양한 처리 능력의 단말기가 함께 이용되고 있다. 이와 같은 다양한 통신망의 구성요소는 통신망 환경을 매우 이질적 (heterogeneous)으로 만들고 있으며[Bane97], 이로 인한 멀티미디어 서비스 사용자의 폭넓은 QoS 요구를 만족시키기 위한 노력이 요구된다.

한편, 계층적 비디오 코딩의 기술을 실시간 멀티미디어 통신에 적용하려는 기존의 연구는 대부분 압축된 비디오 정보를 하나의 기본 계층과 복수 개의 상위 계층으로 나누어 각 계층을 서로 신뢰도가 다른 채널로 전송하려는 시도이며, 이들 연구의 초점은 패킷 손실을 최대한 복구함으로써 패킷 손실에 대한 비디오 데이터의 품질저하를 최소화하려는 방식에 모아지고 있다. 또한, 기존의 연구는 사용자 단말까지의 전송 경로의 대역폭의 다양성이나 사용자 단말 환경의 다양성에 무관하게 모든 계층의 비디오 정보를 최종 사용자 단말까지 전송하고, 사용자 단말에서 자신의 환경에 적합하게 각 계층을 취사 선택하도록 하고 있다. 그러나, 이러한 접근 방법은 불필요한 계층의 비디오 정보를 전송함에 따른 대역폭의 낭비를 초래하고 있으며, 이로 인하여 전송지연율이 높아지게 되는 문제가 있다. 이러한 접근 방식은 다양한 사용자 및 통신망 환경을 동일시하고 동일화된 통신망 등의 환경에서 패킷 손실이나 지연 변이 (jitter) 등의 QoS를 최대한 만족시키려는 획일적 통신망 패러다임 (Homogeneous networking)으로 분류될 수 있다. <그림 1>은 획일적 통신망 패러다임의 서비스 개념도를 보이고 있다.

그러나 통신망 환경 등의 다양성을 실제적으로 수용할 수 있는 이질적 통신망 패러다임 (Heterogeneous networking)의 필요성이 부각되고 있으며, 멀티미디어 통신이나 멀티미디어 처리 분야에서 많은 연구가 도출되고 있는 상황이다.

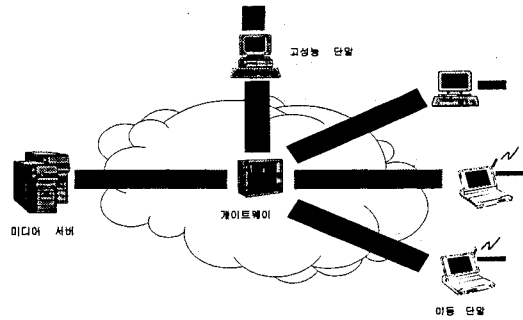


그림 1. 획일적 통신망 패러다임에서의 멀티미디어 서비스 개념도

<그림 2>에서 이질적 통신망 패러다임에서의 멀티미디어 서비스 개념도를 보이고 있으며, 각 사용자의 환경에 따라 차등적인 멀티미디어 서비스를 제공하는 형태를 나타내고 있다.

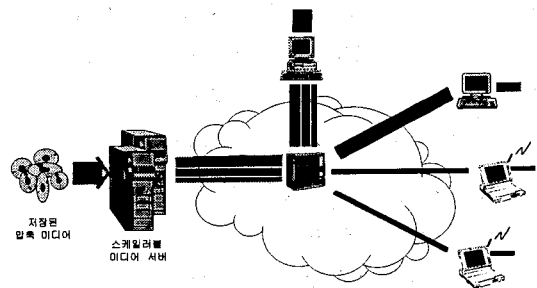


그림 2. 이질적 통신망 패러다임에서의 멀티미디어 서비스 개념도

이질적 통신망 패러다임을 지원하기 위한 기반 기술로서 현재 연구 논점이 되고 있는 분야는 다음과 같다.

- 압축 비디오의 스케일러블 전송을 위한 압축 데이터의 세분화 및 복원 기술
 - 상이한 QoS 매핑을 위한 게이트웨이 중계 기술
 - 서로 다른 신뢰도를 가지는 통신 채널을 통한 각 세분화된 데이터의 전송 기술
 - 패킷 손실에 대한 각 채널 계층별 보정 기술
- 이 중에서 통신망을 통하여 실제 전송되는 데

이타 자체가 스케일러블 전송에 적합한 형태로 표현되어 있어야하기 때문에 압축 비디오 데이터의 세분화 처리 및 복원 처리 기술에 대한 기반 연구가 선행되어 한다.

3. 멀티포인트 통신과 미디어 스케일링

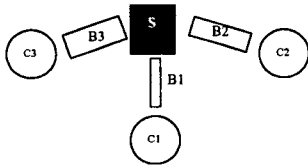
이질적인 통신망 환경에서 사용자들의 다양한 QoS를 만족시키기 위한 연구가 활발히 진행되어 왔으며, 다음과 같이 두 부류로 구분할 수 있다. 한 그룹은 QoS 요구사항을 채널의 주어진 용량에 맞추기 위하여 대역폭 할당이나 허용 제어, 자원 예약 등의 기법을 연구하고 있다 [Chan98, Krun97]. 또 다른 그룹은 미디어 스케일링 기법을 통하여 멀티미디어 데이터의 데이터량을 조절함으로써 채널에 실리는 부하를 경감시키려는 노력을 하고 있다 [Delg94, Zeng96, Camp98]. 이 두 가지 접근 방식은 각각 트랜스포트 시스템 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 연구는 이 중에서 후자의 경우인 미디어 스케일링 접근방식을 따르고 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 데이터를 조정하는 것이다. MPEG-2 비디오의 미디어 스케일링에 이용되는 대표적인 방법으로는 프레임율 조정 (frame rate control)과 충실도 조정

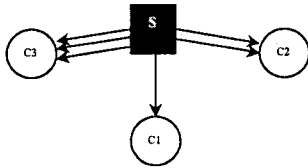
(fidelity control)을 들 수 있다. 프레임율 조정은 수신 측에서 비트스트림의 디코딩 시의 중요도가 낮은 프레임을 생략하고 전송하는 방식이다. 충실도 조정은 DCT 계수 중에서 중요도가 낮은 계수인 AC항의 일부를 생략하고 전송하는 방식이다. 영상회의처럼 데이터가 실시간으로 발생하는 서비스에서는 비디오 인코딩의 입력 변수인 양자화 계수를 변화시키거나 발생 프레임율을 인코더에서 제어함으로써 원하는 QoS에 대한 데이터를 인코더로부터 얻을 수 있다. 그러나, VOD (Video On Demand)나 하이퍼미디어 등과 같은 저장형 서비스에서는 이미 코딩되어 저장된 MPEG-2 비트스트림을 다루게 되므로, 데이터율의 조정을 위하여 인코딩시의 매개변수를 조절하는 방식은 불가능하다. 따라서, 저장형 서비스를 위해서는 별도의 미디어 스케일링 기법이 필요하다.

이미 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 있게되며, 이는 곧 미디어 서버의 불필요한 저장 공간 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 3>은 이와 같은 직관적인 미디어 스케일링 방식의 한계를 설명하고 있다.

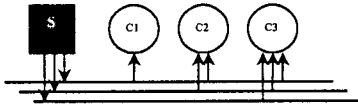
<그림 3>-(a)는 미디어 서버 S와 클라이언트



(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 3. 다양한 대역폭의 이질적인 통신망 환경

C1, C2 및 C3가 이질적인 통신망 환경에서 연결되어 있는 상황을 보이고 있다. 이때, 서버와 클라이언트 C1, C2 및 C3를 연결하는 각 링크의 대역폭을 각각 B1, B2, 그리고 B3라 했을 때, 각 대역폭의 관계는 $B_3 > B_2 > B_1$ 라고 가정한다. 즉, C3는 가장 넓은 대역폭의 링크로 연결되어 있고, C1은 가장 좁은 대역폭의 링크로 연결된 경우이다. <그림 3>-(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 3>-(c)의 경우처럼 모든 클라이언트가 동일한 멀티캐스트 세션에 참여하고 있는 경우에는 여러 버전의 비트스트림을 유지하는 직관적인 방식은 더욱 한계를 가지게 된다. 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 낭비하게 되는 링크가 발생한다.

4. MPEG-2 비디오 코딩

4.1 MPEG-2 비디오 코딩의 원리

MPEG-2 비디오의 코딩 원리는 이동 보상 코딩과 블럭 기반 변환 코딩을 근간으로 한다 [Mpeg94]. MPEG-2 비디오 비트스트림은 <그림 4>에 보인 바와 같이 다단계의 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 8×8 픽셀 배열이며 변환 코딩(transform coding)의 단위로 이용된다. 매크로 블럭은 이동 보상 코딩의 단위로 이용되며, 2×2 개의 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

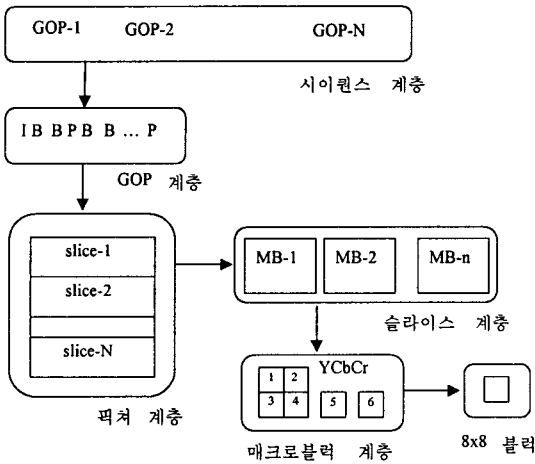


그림 4. MPEG-2 비디오의 다단계 문법 구조

픽처 내에서 수평방향으로 연속된 매크로 블럭들은 하나의 슬라이스를 구성하며, 슬라이스는 데이터 손실의 파장을 줄이기 위한 재동기 단위로 사용된다.

각 8×8 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 스트림으로 재배열되며, 이는 다시 Run-Length 코딩 (RLC)을 거치게된다. 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. <그림 5>는 이와 같은 이동보상 코딩에 의한 각 타입의 픽처간의 종속성을 보이고 있다. 이와 같은 픽처간의 종속성이 의미하는 바는 I-픽처는 독자적으로 디코딩이 될 수 있는 반면, P-픽처나 B-픽처는 다른 참조 픽처가 존재해야만 디코딩이 될 수 있다는 것이다. 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.

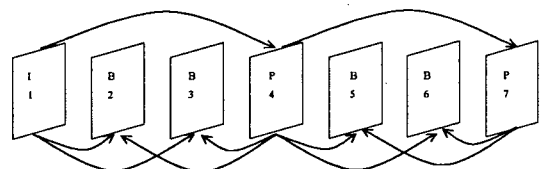


그림 5. 픽처 타입간의 데이터 종속성

코딩된 각 픽처의 크기를 비교해보면, 일반적으로 I-픽처의 크기는 P-픽처보다 크고, P-픽처의 크기는 다시 B-픽처의 크기보다 큰 관계가 성립된다. 크기의 비율은 비디오 시퀀스의 각 영상의 복잡도 및 움직임의 속도 등의 특성마다 차이가 있다. <그림 6>은 IBBPBBPBBP의 형태로 코딩된 Flower Garden 시퀀스에 대하여 각 픽처의 크기를 비교한 내용이다. 이 Flower Garden 시퀀스의 경우에는 I-픽처, P-픽처, B-픽처의 크기 비율이 대략 7:3:1 정도를 보이고 있다. 한편 I-픽처, P-픽처, B-픽처의 개수의 비율은 1:3:6으로 코딩된 것이기 때문에, 각 타입별 픽처의 데이터 양은 대략 전체적으로 비슷하게 이루어진다. 픽처 타입간 데이터 양의 이러한 분포 특성으로부터 데이터 종속성을 고려하여 다른 픽처의 복원에 영향을 주지않는 픽처의 생략을 통한 데이터 양 조정이 가능해진다.

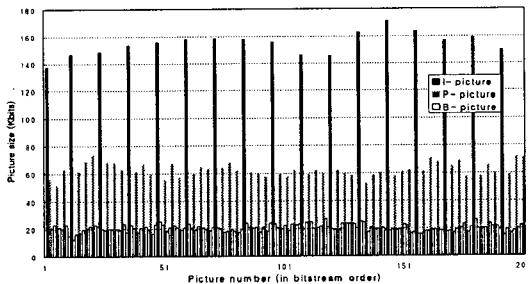


그림 6. I 타입, P 타입 및 B 타입 픽처들의 데이터 양 분포 : 1.1 Mbps로 MPEG-2 코딩된 Flower Garden 비디오 시퀀스

4.2 MPEG-2 비디오의 비트스트림 스케일러빌리티

MPEG-2 비디오에서는 계층적 코딩을 위한 문법 구조가 정의되어 있으며, 이를 스케일러빌리티 확장(scalability extension)이라고 부른다. MPEG-2 비디오의 스케일러빌리티 확장은 공간

스케일러빌리티(spatial scalability), 시간 스케일러빌리티(temporal scalability), SNR 스케일러빌리티(SNR scalability) 및 데이터 분할(data partitioning)으로 구분된다.

가) 공간 스케일러빌리티

공간 스케일러빌리티는 공간 해상도의 확장성을 제공하기 위한 계층적 코딩 개념이다. 공간 해상도가 낮은 계층을 기본 계층(basement layer)라 하고, 공간 해상도가 높은 계층을 상위 계층(enhancement layer)라 한다. 기본 계층은 독자적인 형태로 코딩이 되며 최소한의 공간 해상도를 제공한다. 상위 계층은 기본 계층의 영상을 업 샘플링하고 공간 인터플레이션한 결과와 원래 영상과의 차이를 코딩한 것이며, 이는 최대 공간 해상도를 제공한다. 각 계층은 서로 다른 프레임 크기, 서로 다른 프레임율, 혹은 서로 다른 색상 포맷을 가질 수 있다. 공간 스케일러빌리티의 전형적인 응용을 보면, 기본 계층을 SIF¹⁾ 해상도의 MPEG-1 호환 비트스트림으로 형성하고, 상위 계층은 CCIR 601²⁾ 해상도의 비디오를 형성하는데 사용될 수 있다.

나) 시간 스케일러빌리티

공간 해상도는 같지만 1초당 프레임율이 다른 것들을 효율적으로 보낼 때 시간 스케일러빌리티를 사용한다. 시간 해상도가 낮은 계층을 기본 계층이라 하고, 시간 해상도가 높은 계층을 상위 계층이라 한다. 상위 계층은 기본 계층에 이용한 시간적 예측을 통하여 코딩되며, 기본 계층은 통상

1) SIF은 Source Input Format의 약자이며 360×240 픽셀 (30 Hz) 혹은 360×288 픽셀 (25 Hz)의 해상도를 가진다. 색상 포맷은 4:2:0을 채택한다.
2) CCIR 601은 디지털 TV를 위한 표준이며 720×480 픽셀 (30 Hz) 및 720×576 픽셀 (25Hz)의 밝기 해상도를 가진다. 색상 포맷은 4:2:2를 채택한다.

적인 MPEG-2 비디오 코딩방식으로 부호화된다. 각 계층은 프레임 크기와 색상 포맷은 모두 같고, 프레임 울만 달라야한다.

다) SNR 스케일러빌리티

SNR³⁾ 스케일러빌리티는 여러 단계의 화질로 비디오를 전송하고자 하는 응용에서 주로 이용된다. 각 계층은 동일한 공간 해상도와 시간 해상도를 가져야하고, 화질만 다를 수 있다. 기본 계층은 독자적인 형태로 코딩되며 최저 비디오 화질을 결정한다. 상위 계층은 기본 계층의 DCT 계수의 양자화 스텝을 세밀하게 조정하여 얻어진 결과를 코딩한 것이며, 최대 화질을 결정한다. SNR 스케일러빌리티는 각 계층을 서로 다른 에러율을 가지는 채널로 전송함으로써 전송 에러에 대하여 향상된 대응력을 제공한다.

라) 데이터 분할

데이터 분할을 이용하면 하나의 비트스트림을 파티션(partition)이라 불리는 두 개의 계층으로 나눌 수 있다. 예를 들어, 첫 번째 파티션은 헤더 정보나 이동 벡터 등의 중요한 헤더 정보를 포함하도록 하고, 두 번째 파티션은 나머지 비트스트림을 포함하도록 할 수 있다. 이 방식은 하나의 비트스트림에 대해서 두 개의 채널을 할당할 수 있는 응용에서 활용 가능하다.

4.3 MPEG-2 비디오 스케일러빌리티의 한계

MPEG-2는 다양한 형태의 서비스를 제공하기 위하여 위에서 기술한 바와 같은 스케일러빌리티 모드를 제공하고 있으나, 이질적 통신망 환경에 적용하기에는 다음과 같은 한계가 있다.

- 각 스케일러빌리티 모드에서는 두 가지 계층만을 정의할 수 있다. 따라서, 폭넓은 비트율의 요구를 만족시킬 수 있는 다양성이 부족하다. 이는 비록 두가지 이상의 스케일러빌리티 모드를 혼합하여 적용하게 되는 경우에도 혼용될 수 있는 폭이 제한되기 때문에 동일한 문제가 발생할 수 있다.
- 일단 코딩이 완료된 후에는 사용자의 QoS 요구에 대한 적절한 비트율의 적응형 서비스를 제공하기 불가능하다. MPEG-2의 스케일러빌리티 모드의 목표는 하나의 비트스트림을 전송하여 사용자의 단말 환경에 따라 표준 디지털 TV 혹은 HDTV 등의 품질 등의 서로 다른 품질을 제공하는 것이었기 때문이다.
- MPEG-2의 기본 계층으로 가능한 최하위 규격은 MPEG-1의 규격이다. 따라서, 상위 계층을 제외하여도 기본 계층만의 데이터 양도 이미 방대하여 대역폭이 작은 채널에 대한 서비스의 열악한 품질에 대해서는 이미 지적되고 있는 사항이다.
- 따라서, MPEG-2 비디오의 기본 계층에 대하여 미디어 스케일링을 통한 세분화를 하고, 각 세분화된 데이터를 사용자의 QoS 요구에 적응하여 전송할 수 있는 체계가 필요하다.

5. 스케일러블 미디어 객체 프레임워크

본 절에서는 QoS 적응형 멀티미디어 통신 서비스를 위한 MPEG-2 비디오의 미디어 스케일링 및 표현 구조인 스케일러블 미디어 객체 프레임워크[Kim99]를 소개한다. 하나의 MPEG-2 비디오 비트스트림은 세부 비트스트림의 집합으로 분할되며, 이러한 세부 비트스트림을 스케일러블 미디어 객체(scalable media object)라 한다. 각 스케일러블 미디어 객체들은 MPEG-2 비디오 문

3) SNR은 Signal to Noise Ratio(신호대 잡음비)의 약자이며 주어진 영상의 화질을 비교하는 기준으로 이용된다.

법에 호환되는 비트스트림으로 복원되기 위하여 서로 의미있는 관계설정을 통하여 재결합되어야 하는데, 이러한 관계설정을 표현하는 데이터 모델을 스케일러블 미디어 객체 프레임웍이라 한다. 이때, 스케일러블 미디어 객체간의 관계설정은 MPEG-2 호환 표현이 되기 위하여 필요한 데이터 종속성을 기반으로 형성된다. 주어진 MPEG-2 비디오 비트스트림을 스케일러블 미디어 객체로 분할하는 기법으로서 프레임 울에 따른 시간적 스케일링과 화질에 따른 충실도 스케일링을 적용한다.

5.1 미디어 스케일링

가) 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 프레임을 삭제하는 형태의 단순한 프레임 울 조정 방식은 적용이 불가능하다. 따라서, 픽처 타입에 따른 프레임 울 조정 방식이 필요하다.

시간적 스케일링은 MPEG-2 비디오 비트스트림으로부터 픽처 계층을 픽처 타입에 따라 분할하고, 이들을 다시 MPEG-2 비디오 비트스트림으로 복원하는 과정을 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. I-서브스트림은 MPEG-2 비디오의 선택소와의 호환성을 가지며, 독자적으로 디코딩 될 수 있다. P-서브스트림과 B-서브스트림

은 상위 계층 헤더 정보를 가지고 있지 않고, 이동 보상 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.

나) 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.

8×8 픽셀의 2차원 블럭을 변환한 64개의 2차원 DCT 계수는 지그재그 스캐닝이나 교체(alternate) 스캐닝 과정을 통하여 1차원 계수열로 변환된다. 64개의 1차원 DCT 계수는 N 개의 계수 그룹 G ,로 분할되며 이 계수 그룹들은 순서 있는 집합 (ordered set)인 S 에 포함된다. 이때, N 값의 범위는 $0 < N \leq 64$ 이며, i 값의 범위는 $1 \leq i \leq N$ 이다. 순서 있는 집합 S 의 멤버의 갯수인 N 은 DCT 계수 집합 S 가 표현할 수 있는 충실도의 계층 수를 나타낸다. S 의 멤버인 계수 그룹 G_i 는 k 개의 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. <그림 7>은 2차원 DCT 계수를 1차원으로 변환하여 계수 집합 S 를 형성하는 충실도 스케일링 분할 과정을 보이고 있으며, N 과 k 가 서로 다른 두 가지 분할

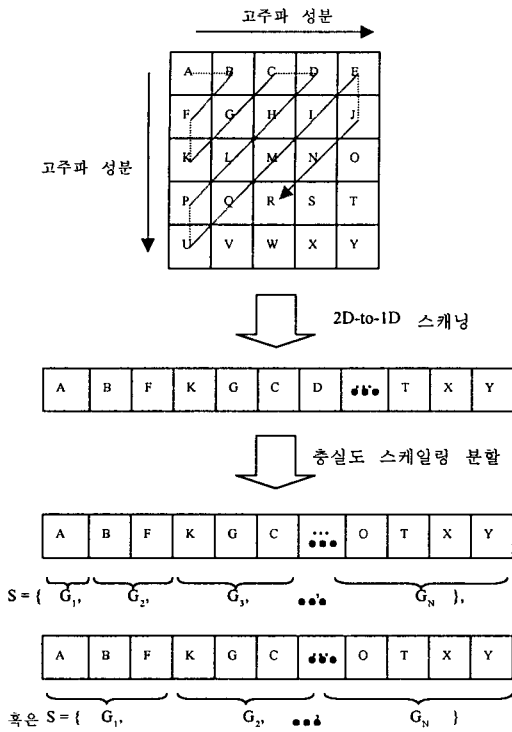


그림 7. 총실도 스케일링에 의한 각 블록의 DCT 계수 분할

방법을 나타내고 있다.

$i < j$ 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_j 는 그룹 G_i 에 종속된다고 한다. 그리고, 한 블록에서의 최저주파향인 DC 항은 항상 G_1 그룹에 속하며, G_1 그룹의 첫 번째 요소이다. 즉, 그룹 G_j 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 변환하는

데에도 적용되기 때문이다. 따라서, 총실도 스케일링은 시간적 스케일링의 결과인 I-서브스트림, P-서브스트림 및 B-서브스트림에 모두 적용될 수 있다. 이러한 특성을 기반으로 하여 시간적 스케일링과 총실도 스케일링이 복합적으로 적용될 수 있게되며, 이로 인하여 폭넓은 미디어 품질을 제공할 수 있게된다.

5.2 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 있게된다. 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 프레임워크라 부른다.

<그림 8>은 스케일러블 미디어 객체 프레임워크를 보이고 있다. 각 노드 $t_i f_j$ 는 시간적 스케일링의 i -서브스트림과 충실도 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. 특히, <그림 8>은 충실도 스케일링을 적용하여 5개의 계수 그룹만 형성한 경우 ($1 \leq j \leq N, N=5$)를 보이고 있다. G_j 는 임의의 개수의 연속된 1차원 DCT 계수를 가질 수 있다.

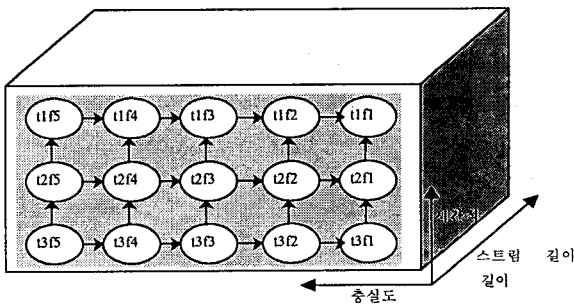


그림 8. 스케일러블 미디어 객체 프레임워크 (노드는 스케일러블 미디어 객체, 링크는 객체들간의 종속성을 표현)

노드 $t_1 f_1$ 은 DC항의 계수만으로 하나의 블록이 표현된 I타입의 픽처와 상위 계층 헤더로 구성된 비트스트림을 의미한다. 다시 말하면, $t_1 f_1$ 은 $t_1 f_F$ 으로 표현되는 I-서브스트림에 충실도 스케일링을 적용하여 DC항의 값만으로 각 블록을 대표하도록 하는 (이를 $t_T f_1$ 으로 표현함) 비트스트림이다. 이때, I-서브스트림은 독자적 표현형태⁴⁾ (self-contained)이고, DC항도 역시 독

자적 표현형태이기 때문에 $t_1 f_1$ 이 독자적 표현 형태임은 명확하다. 이 사실은 노드 $t_1 f_1$ 이 스케일러블 미디어 객체 프레임워크에서 다른 스케일러블 미디어 객체가 디코딩되는 과정에서 가장 필수적으로 필요한 기본 객체임을 의미한다.

<그림 8>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 $T2F2$ 라 부른다. 따라서, 특정 비트율을 가지는 MPEG-2 비디오 호환 비트스트림을 구성하기 위해서는 비트율을 만족시키는 스케일러블 미디어 객체 그룹 $TiFj$ 가 필요하다. 그리고 $TiFj$ 를 형성하기 위해서는 $i * j$ 개의 스케일러블 미디어 객체 $t_m f_n$ ($m \leq i, n \leq j$)들이 필요하다.

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, 디코드되는 과정을 거치게 된다. 클라이언트가 요구한 QoS 값을 비트율로 환산하고, 환산된 비트율에 해당되는 스케일러블 미디어 객체 그룹을 선정한다. 특정 스케일러블 미디어 객체 그룹이 선정되면, 그에 포함되는 스케일러블 미디어

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

어 객체를 전송함으로써 클라이언트의 QoS 요구를 만족시킬 수 있게 된다.

이에 따라, 클라이언트들이 이질적 통신망 환경에서 어떤 형태로 미디어 서버에 연결되어 있는지 무관하게 클라이언트로부터 발생하는 다양한 QoS 요구에 적응할 수 있는 기반을 스케일러블 미디어 객체 프레임워크가 제공한다. 앞의 3절의 <그림 3>-(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. 직관적인 미디어 스케일링 방식에서도 각 QoS를 만족시킬 수는 있었으나, 중복된 데이터에 따른 저장장치의 낭비가 발생하였다. 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 3>-(c)에서는 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.

6. 맺음말

본 고에서는 멀티미디어 통신 서비스를 구현할

때에 간과할 수 없는 통신망 환경의 이질성을 지적하였다. 이러한 이질적 통신망 환경에서 서로 다른 QoS를 만족시키기 위해서는 주어진 대역폭을 분할하고 예약하는 등의 전송 계층에서의 노력만으로는 문제 해결의 범위가 제한되어 있음을 논하였다. 이의 해결을 위하여 전송 계층의 노력과 더불어 응용 계층의 핵심 부분인 멀티미디어 데이터 자체를 스케일링함으로써 다양한 QoS에 적극적으로 적응할 수 있는 접근 방법의 중요성을 제기하였다.

또한 QoS 적응형 멀티미디어 서비스를 위한 MPEG-2 비디오의 미디어 스케일링 및 표현구조의 사례연구로서 스케일러블 미디어 객체 프레임워크[Kim99]를 소개하였다. 본 고에서는 지면 관계상 생략되었으나 스케일러블 미디어 객체의 특성이 데이터 량의 다양성 측면, 미디어 스케일링에 따른 신호왜곡 특성 측면, 그리고 신호왜곡의 주관적인 평가를 위한 복원화질 비교 측면에서 분석되었으며, 분석결과는 [Kim99]를 참고하기 바란다.

참 고 문 헌

- [Bane97] A. Banerjea, "Heterogeneous networking", IEEE Multimedia, Vol. 4, No. 2, pp. 84-87, 1997
- [Beck98] C. J. Beckmann, A. A. Moin, S. Nog, "Bandwidth reservation with selectable bit-rate streams", ACM Multimedia Systems, Vol. 6, pp. 219-231, 1998
- [Belz94] B. Belzer, J. Liao, J. D. Villasenor, "Adaptive video coding for mobile wireless networks", IEEE Int'l Conf. on Image Processing (ICIP'94), Vol. 2, pp. 972-976, Oct. 1994
- [Camp98] A. T. Campbell, G. Coulson, D. Hutchison, "Transporting QoS adaptive flows", ACM Multimedia Systems, Vol. 6, pp. 167-178, 1998
- [Chan98] D. L. S. Chan, S. T. Chanson, "Scalability

- support for multiparty multimedia communications", ACM Multimedia Systems, Vol. 6, pp.75-87, 1998
- [Delg94] L. Delgrossi, C. Halstrick, D. Hehmann, et. al., "Media scaling in a multimedia communication system", ACM Multimedia Systems, Vol. 2, pp. 172-180, 1994
- [Ferg98] P. Ferguson, G. Huston, Quality of Service : Delivering QoS on the Internet and in Corporate Networks, Wiley Computer Publishing, 1998
- [H261] ITU-Telecommunication H.261, Video CODEC for audiovisual services at px64 Kbits/s, 1995
- [Han97] J. K. Han, G. C. Polyzos, "Multi-resolution layered coding for real-time image transmission: architectural and error control considerations", Real Time Imaging, Academic Press, 1997, <http://www-csl.ucsd.edu/>
- [Kim99] H. C. Kim, Y. J. Won, C. G. Jeong, "Scalable media object framework for heterogeneous networking environment", IEEE International Conf. on Telecommunications (ICT'99), June 1999 (to appear).
- [Krun97] M. Krunz, S. K. Tripathi, "Impact of video scheduling on bandwidth allocation for multiplexed MPEG streams", ACM Multimedia Systems, Vol. 5, pp. 347-357, 1997
- [Mpeg93] ISO/IEC JTC1/SC29/WG11, Coding of moving pictures and associated audio for digital storage media at up to about 1.5 Mbit/s - Part 2: Video, IOS/IEC 11172-2, International Standard, 1993
- [Mpeg94] ISO/IEC JTC1/SC29/WG11, Coding of moving pictures and associated audio -- Video part, ISO/IEC 13818-2, Draft International Standard, March 1994
- [Vers98] O. Verscheure, X. Garcia, G. Karlsson, J.P. Hubaux, "User-oriented QoS in packet video delivery", IEEE Network, Vol. 12, No. 6, pp. 12-21, Nov./Dec. 1998
- [Zeng96] 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



김 형 철

- 1986년 인하대학교 전자계산학과(이학사)
 - 1988년 KAIST 전산학과(공학석사)
 - 1995년 KAIST 전산학과(공학박사)
 - 1988년 ~ 현재 ETRI 멀티미디어연구부 선임연구원
 - 관심분야 : video coding, multimedia communication, multimedia collaboration
-
-