

지연에 민감한 대화형 서비스를 위한 동영상 전송을 평활화 연구

장 승 기, 서 덕 영
(Seung Gi Chang, Doug Young Suh)

경희대학교 전자공학과 뉴미디어 연구실
Kyunghee University NewMedia Lab., Dept. of Electronic Engineering

ABSTRACT

This paper¹⁾ describes two traffic shaping methods for VBR(variable bit rate) video bit streams encoded in the MPEG-2 syntax. Difficulties in control of VBR video traffic are lessened by traffic shaping. Burstness of single layer MPEG-2 video can be reduced by performing intra-refresh in more than one consecutive frames. In two layer encoding of spatial scalability, burstness can be reduced by setting the temporal locations of GOP starting frame of a layer, differently from the other. Queueing analysis shows that these two methods outperform conventional encoding schemes in terms of temporal and semantic QoS requirements.

1. 서론

멀티미디어 서비스는 크게 '방송'(broadcasting)과 '대화'(interactive)로 나눌 수 있다. '방송'은 지연에 대하여 덜 민감하여 수 초의 지연이 허용된다. 그러나 실시간 대화형 멀티미디어 서비스는 지연에 민감하여 수십 ms이상 지연이 일어나면 사용자가 불쾌감을 느끼게 된다. 압축된 동영상은 기본적으로 VBR 데이터이다. 피크 비트율이 커질수록 VBR 데이터를 실시간으로 전송하는데 어려움이 커진다. 본 연구는 MPEG-2 인코딩 방식을 적절히 선택하여 동영상 트래픽을 평활화하여 지연을 최소화하는데 그 목적이 있다.

현재 국내외에서 B-ISDN을 위해 ATM

(asynchronous transfer mode) 프로토콜과 관련된 기술이 활발히 개발되고 있다. ATM은 이 B-ISDN에서의 표준으로 음성이나 데이터 뿐만 아니라 음성 및 데이터 트래픽과는 여러면에서 다른 특성을 갖는 영상 데이터를 동시에 다룰수 있는 앞으로의 통신망에서 필요한 핵심 기술이다. 특히 영상 데이터는 데이터량이 클 뿐만 아니라 버스트적이고, 지연에도 민감하다. ATM은 AAL 1만 제외하고 모든 AAL 타입이 VBR 트래픽을 지원한다. 이것이 ATM에서는 다양한 멀티미디어 서비스를 할 수 있는 장점이 된다. VBR 트래픽은 사용자와 네트워크에서 제어할 기술이 필요한데, 이를 각각 'shaping', 'policing'이라고 한다. Shaping은 call-setup시 ATM 교환기와 정한 약속을 지키기 위한 터미널의 노력이다.

그 약속 내용에는 평균 셀 율(MCR; mean cell rate), 최대 셀 율(PCR; peak cell rate), 최대 버스트 크기(MBS; maximum burst size)등이 있

본 연구는 한국통신의 '96정보통신기초연구과제'와 정보통신연구관리단의 '96대학기초연구지원사업' 지원하에 이루어졌음을 밝힙니다.

다. VBR로 인코딩 된다면 영상 화질이 영상의 복잡함에 관계없이 일정하게 유지될 수 있고 인코딩된 데이터는 인코딩과 동시에 전송될 수 있다. 이와는 반대로 CBR로 인코딩된 비디오는 화질면에서 변동을 가지고 있다. 또한 CBR을 하기 위한 버퍼는 지터(jitter)를 유발한다. VBR의 하드웨어는 CBR 보다 훨씬 간단한 반면 트래픽 제어는 더욱 어렵다. 비트율 변동의 측정수단 중 하나는 버스트니스(burstiness)인데 평균 비트율에 대한 최대 비트율의 비로 정의 된다. 본 논문은 지연 발생이나 버퍼 공간 필요 없이 버스트니스를 줄이는 두가지 방식을 제안한다. Yang[2]은 버퍼를 사용함으로써 VBR 비디오 트래픽을 평활화 하는데, 이는 어느 정도의 지연변이를 발생시킨다. 비디오 데이터의 평활화 과정은 MPEG-2 인코더를 거친후에 하게 된다. 하지만, 본 논문에서 제안한 방식은 MPEG-2를 어긋남 없이 MPEG-2 인코더 내에서 평활화 된다. 따라서 원하는 비트 스트림은 기존의 MPEG-2 디코더로 디코딩 된다. 또한 본 논문에서는 계층화 인코딩의 비디오 트래픽을 평활화 하는 방식을 제안한다. MPEG-2는 하나의 비디오 데이터 스트림을 둘 혹은 세 개의 스트림으로 인코딩 할 수 있다. MPEG-2 표준에서 정의된 4가지 종류의 계층화 방식 가운데 여기서 제안한 방식은 공간적 계층화에 해당한다. ATM 망에서 한 서비스는 같은 VP(virtual path)와 다른 VC(virtual channel)을 갖는 다중 비트 스트림으로 분배될 수 있다. 제안한 두 방식 모두 Intra, Predictive, Bidirectional 등 인코딩 모드에 따라서 이미지의 압축률이 상당히 다르다는 것을 근거로 하고 있다. 우리는 Intra 프레임의 비트율을 분배하는 것이 통계적인 특성이 매우 안정하다는 것을 알았다. 2장에서는 QOS성능 분석 수단의 정의들을 설명하고 3장에서는 단일계층 VBR MPEG-2 비디오에 대한 세이핑 방법을 4장에서는 2계층 VBR MPEG-2 비디오에 대한 트래픽 세이핑 방법을 설명한다.

2. QOS 성능 분석

AAL에서는 실시간 VBR 트래픽에 대하여 3가지의 QOS 파라미터를 정하고 있다. 이중에서 CLR(cell loss ratio)은 정보의 투명성(semantic transparency)에 관계가 있고 평균 셀 전송 지연(CTD; cell transfer delay)과 셀 지연 변이(CDV; cell delay variance)는 시간적 투명성(temporal

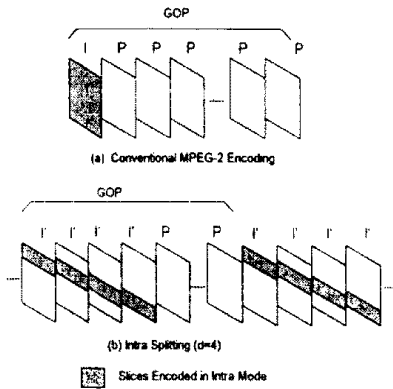
transparency)에 관계가 있다. 버퍼는 지터를 발생시킴으로서 버스트 트래픽을 완화시킬 수 있다. 대체로 광섬유의 SNR이 매우 높기 때문에 ATM 망에서의 셀 손실은 ATM 교환기내에서의 트래픽 제어의 실패에 기인한다고 볼 수 있다. 어떤 셀이 도착하였을 때 버퍼가 채워져 있다면 그 셀은 폐기된다. 특히 원격회의의 비디오 비트스트림과 같은 실시간 트래픽은 요구되는 지연시간 보다 늦게 도착하는 셀에 대하여는 마찬가지로 폐기된다. 우리는 다음과 같이 버퍼의 크기를 정함으로써 두 가지의 셀 폐기 원리를 시뮬레이션할 수 있다.

$$BufferSize = DepartureBandwidthLimit \times Deadline.$$

CLR, CTD, CDV는 서로 밀접한 관련이 있다. 버퍼의 크기와 최대 출력대역폭이 고정 되었다면 CDV와 CTD가 클수록 CLR은 더욱 작아진다. 즉, 시간적 그리고 정보적 QOS 파라미터는 trade-off관계에 있다. 이것들은 서비스의 특성에 따라서 결정이 되는데 예를 들어 방송의 경우는 지연에 크게 민감하지는 않지만, 원격 회의와 같은 경우는 지연이 상당히 중요한 QOS 파라미터가 된다. 본 논문의 분석에서는 교환기에서 최대 지연시간(deadline)을 10ms로 설정한다. 이 시간은 DAVID[5]과 MPEG Real-Time Interface[6]와 같은 문서에 있는 내용을 참고로 하여 정한 것이다. 버퍼는 매 1ms마다 갱신되는 것을 가정한다.

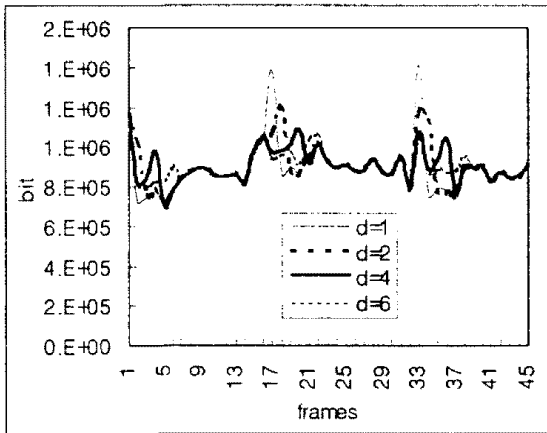
3. Intra 분산 방식

보통 세가지의 인코딩 모드, 즉 Intra, Predictive, Bidirectional의 비트율 비는 8:3:2[7]이다. 이것은 Intra 퍼크를 평활화 할 때는 2-3 프레임 주기의 지연이 요구된다는 것을 의미한다. 본 논문에서는 하나 이상의 연속되는 프레임, 즉 의사 Intra 프레임(PIF; pseudo Intra frame)으로 본래 하나의 Intra 프레임의 Intra 블럭들을 분산하는 방법을 제안한다. Intra 분산을 4개의 프레임(d=4)으로 분산한다면 그림 1(b)에서 보여주는 것과 같이 4개의 프레임 주기 동안 4개의 PIF 각각마다 전체 macroblock의 1/4만큼씩 인코딩 된다. 완전한 Intra 분산을 위해서 약간의 overhead가 필요하다.



[그림 1] Intra 분산 기법

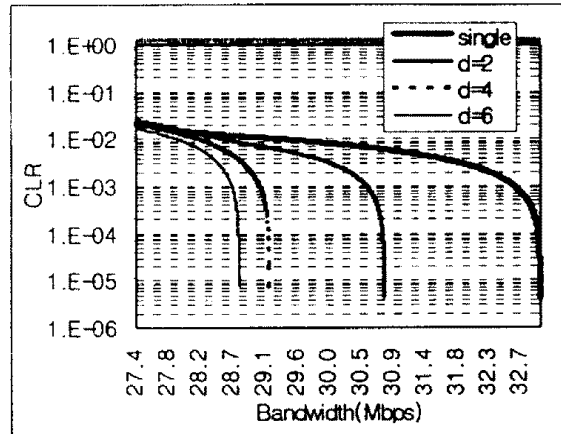
PIF 1에서 1~8 슬라이스가 intra 모드로서 인코딩된 후에 PIF 2의 9~16 슬라이스가 Intra 모드로서 인코딩 된다. PIF 2의 8번째 슬라이스에 있는 Predictive 모드의 매크로 블록이 PIF 1의 9번째 슬라이스를 참조한다고 하면 이 때는 Intra refresh에 의한 error 전파 차단 효과를 보장할 수 없다. 그런 경우에 매크로 블록은 Intra 모드로서 강제로 인코딩 되도록 한다. 본 실험에서는 강제로 Intra 모드로 인코딩 되어야 하는 원하지 않는 Intra 매크로 블록의 수가 예상했던 것보다 훨씬 작음을 보여 주었다. 그림 2는 프레임 비트율이 Intra 분산에 의해 평활화가 잘 됨을 보여주고 있다. PCR(peak cell rate)은 약 25%정도 줄어든 반면, MCR(mean cell rate)의 증가는 무시할 정도이다.



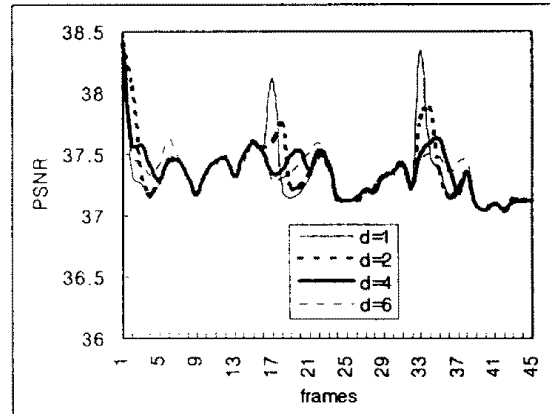
[그림 2] Intra 분산 기법의 프레임 비트율 (Football, GOP=16, M=1, Q=4)

평활화된 비트 스트림을 2장에서 설명한 큐에 적용할 때는 그림 3과 같이 같은 CLR에 대하여 요구되는 대역폭이 12% 정도 감소하였다. 이 그림은 역시 4 PIF 이상으로 Intra를 분산할 때는 이

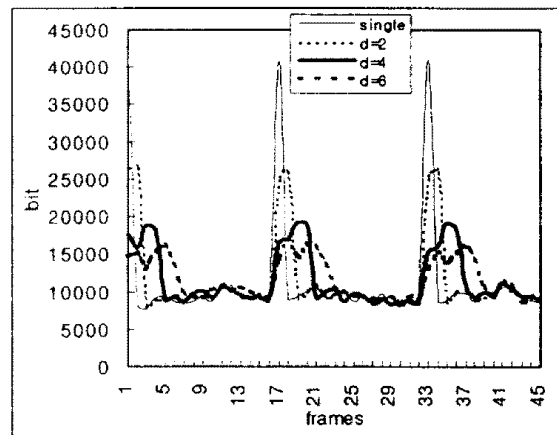
득이 그리 크지 않음을 보여주고 있다. 하나의 또 다른 효과는 프레임 단위의 PSNR의 변동이 그림 4와 같이 줄어든다는 것이다.



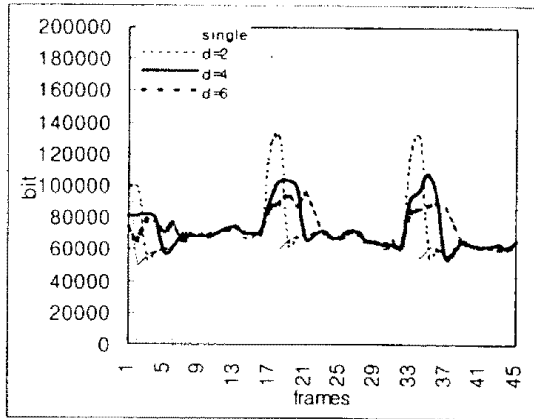
[그림 3] Intra 분산 기법에서의 CLR (Football, GOP=16, Q=4)



[그림 4] Intra 분산 기법의 PSNR (Football, GOP=16, M=1, Q=4)



[그림 5] Intra 분산 기법의 프레임 비트율 (Salesman, GOP=16, M=1, Q=32)

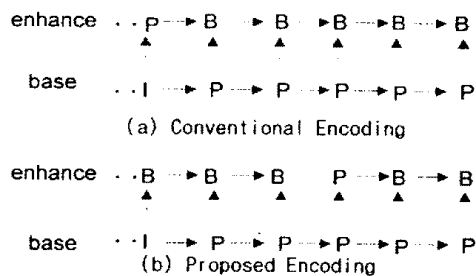


[그림 6] Intra 분산 기법의 프레임 비트율 (Football, GOP=16, M=1, Q=32)

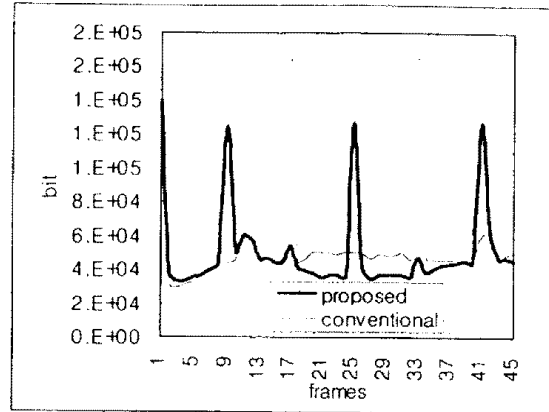
그림 5 및 그림 6은 양자화 크기가 32로서 앞에서의 4보다 큰 값에 대한 실험 결과이다. 그림 5는 Salesman, 그림 6은 Football 영상에 대한 것인데, 그림 2와 비교하여 다른점이 있다. 즉, 양자화 크기 값이 32인 경우가 4인 경우보다 평활화 정도의 효과가 큼을 알 수 있었다. 따라서, 본 실험에서의 결론은 양자화 크기 값이 클수록 평활화 효과가 큼을 나타내고 있다.

4. 계층화 트래픽 셰이핑

MPEG-2에서 정의된 4가지의 계층화 종류 중에서 본 논문에서 제안한 방식은 공간적 계층화, SNR 계층화 방법인데, 이 방법은 유럽에서의 디지털 HDTV 표준이 되고 있다. 이것은 HDTV 및 기존의 TV 모니터의 호환성을 제공하는데에 유용하다. 기본 계층의 Intra 프레임과 고급 계층의 Predictive 프레임은 그림 7(a)에서와 같이 이전 인코딩 알고리즘에서는 같은 프레임에 위치하게 된다. 이와 같은 방법은 해당 프레임 주기에



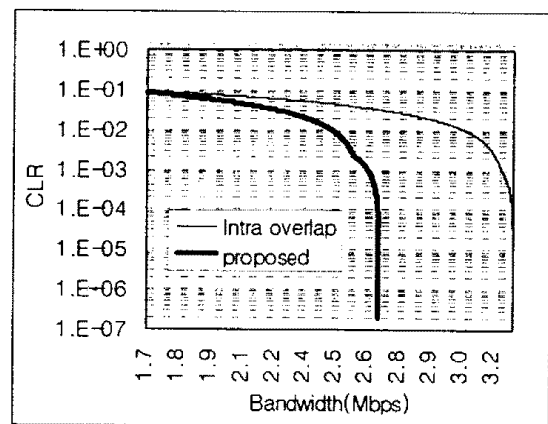
[그림 7] 2계층 평활화 인코딩 방식



[그림 8] 2계층 평활화에서의 프레임 비트율 (기본계층 Q=16, 고급계층 Q=4)

높은 비트율을 유발시켜 다중화기에 영향을 주게 된다. 따라서 본 논문에서는 그림 7(b)와 같은 기본 계층의 Intra 프레임과 고급 계층의 Predictive 프레임이 서로 다른 위치에 놓일 수 있는 인코딩 방법을 제안한다. 이것은 그림 8과 같이 비트 스트림의 전체적인 비트율의 변동을 줄일 수 있다.

이 인코딩 방식의 성능은 기본 계층과 고급 계층의 평균 셀 도착률의 비에 의존한다. 공간적 계층의 인코딩에서 이 비율은 기본 계층의 양자화 레벨에 의해 조절될 수 있고 그 비율이 크면 클수록 더 좋은 성능이 나타남을 알 수 있었다. 아래 그림 9는 기존의 공간적 계층화에서 요구되던 대역폭에 비해 약 20%정도 대역폭이 감소됨을 알 수 있다. 고급 계층에서의 refresh 프레임의 위치는 기본 계층에서와는 독립적으로 결정될 수 있다. 반면, Intra 분산 방법의 PIF는 몇 개의 프레임에 걸쳐 연속적으로 놓이지게 되며 따라서 그 프레임동안 버스트를 초래한다.



[그림 9] 2계층 평활화에서의 CLR (Salesman, GOP=16, Q=4)

5. 결론

본 논문에서는 VBR MPEG-2 비디오 트래픽을 평활화하는 두 가지의 방법을 제안했다. 트래픽이 MPEG 인코더 내에서 평활화되므로 추가적인 지연을 초래하지 않으며, 이 방법은 패킷 교환망이나 ATM 망에서 버스트니스가 큰 동영상 데이터를 실시간으로 전송하는데 매우 효과적인 방식이다. 본 논문의 결과로 대역폭이나 버퍼의 크기와 같은 통신 자원을 절약 할 수 있음을 보였다. 즉, 이것은 같은 물리적인 채널에 통하여, 보다 많은 수의 프로그램을 전송할 수 있음을 의미한다. 또한 대역폭과 셀손실의 관계에서는 버퍼의 허용지연에 따라서 그 상황이 달라지게 된다. 동영상 전송시 지연은 인코더/디코더, 통신망에 대한 지연의 합으로 표시되는데 본 연구에서는 사용자 터미널과 교환기 사이의 지연을 10 ms일 때에 대한 것이다. 아울러 본 논문에서는 약 100 프레임 정도의 영상 데이터에 실험했으나 앞으로 더 긴 시간의 영상 시퀀스에 대한 실험이 필요하다.

6. 참고문헌

- [1] Martin de Prycker, "Asynchronous Transfer Mode", Ellis Horwood, 1991.
- [2] Tao Yang, et. al. "Cell Scheduling and Bandwidth Allocation for Heterogeneous VBR Video Conferencing Traffic", GLOBECOM 95, Vol. 1, pp 371-377, Nov. 1995.
- [3] N. M. Marafih, et. al., "Modeling and Queueing Analysis of VBR Coded Video in ATM Network", IEEE JSAC, Vol. 12, NO. 2, pp121-128, Apr. 1994.
- [4] Ralf Schaefer, "The Status of HDTV in Europe", IEEE Communications Magazine, pp120-125, June, 1996.
- [5] DAVIC 1.0 Specification, Revision 5.0 Berlin Draft, Part 1. Description of DAVIC Functionalities, December, 1995.
- [6] ISO/IEC 13819-9, MPEG, Real-Time Interface Specification, International Standard, N1273, July 1996.
- [7] ISO/IEC/JTC1/SC29/WG11, AVC-445, Test Model 4, Chapter 10, January, 1993.