

효율적인 Scalable Video Streaming을 위한 Cross Layer Optimization

Cross Layer Optimization for Scalable Video Streaming

윤민영, 조희영, 서덕영
Min-young Yoon, Hee-young Cho, Doug Young Suh

Abstract - As further Studies on 4th generation mobile telecommunication are progressed, the importance of a Cross-Layer is being increased. However, it has focused on coordination model only between MAC layer and PHY layer. It is necessary to expand into IP layer and upper layers. In this paper, we introduce a Cross-layer optimization which can be used to transmit video data with effect by managing resources among layers. It can gives further more adaptive method to solve QoS model problem than single layer.

Key Words : Scalable Video Streaming , Optimization, FEC

1. 서론

최근 Home Entertainment 및 Home Networking 시장의 급속한 성장과 함께 Wired 및 Wireless A/V streaming 과 관련된 기술 및 상품의 개발에 학계 및 산업계의 활동이 두드러진 분야로 대두되어왔다. 네트워크를 통해서 비디오 스트리밍을 하는 시스템은 수년 전에 개발되었고 이제 그 서비스가 보편화 되고 있다. 특히, 무선 인터넷망에서는 패킷 당 요금을 계산하므로 최소의 데이터를 전송하면서 최대의 품질을 보장하는 것은 매우 중요한 issue이다. 유선망에서도 IPv6 를 도입하고 QoS (Quality of Service)를 차별화하는 프로토콜의 도입이 조만 간에 이루어질 것이다.

네트워크를 통하여 전송되어야 하는 미디어 중 가장 제약 조건이 많은 비디오의 압축 방식에서도 다양한 논의가 진행되고 있다. SVC(Scalable video coding)는 90년대 중반에 표준화된 MPEG(Moving Picture Expert Group)-2에 포함되어 있었으나 이용되지 않고 있었다. 그러나 최근에 채널상황이 가변적인 인터넷이나 무선망을 통하여 비디오를 전송되면서 SVC에 관심이 높아지고 있다. MPEG-4, H.264 표준에서는 예러에 대한 강인성을 부여하는 방법과 시간에 따라 시시각각으로 변화하는 네트워크의 품질에 적응하도록 하나의 비디오를 품질에 따라 다중의 계층(layer)으로 여러 개의 비트스트림으로 부호화하는 계층부호화(scalable encoding) 및 FGS (Fine Granular Scalability)에 대하여 많은 연구가 계속 진행되고 있다.[2][3][4] 또한, MPEG-2에서는 혼재망 (heterogeneous networks)에서의 비디오 스트리밍을 위해 더 효율적인 SVC 방식을 표준화 하고 있다. 이러한 비디오 혹은 멀티미디어 서비스의 품질 향상을 위해서는 네트워크의 각 계층에서 유기적으로 이루어져야 한다. Cross Layer는 layer 간의 상호 협력이라는 새로운 모델을 제시하며 기존 시스템의 성능을 획기적으로 향상 시킬 수 있는 가능성을 내세우고 있다. 본 논문에서는 SVC coding을 각 Layer마다 차별적인 packet size와 bit Energy를 이용한 Cross Layer를 제안하고 있다. 이는

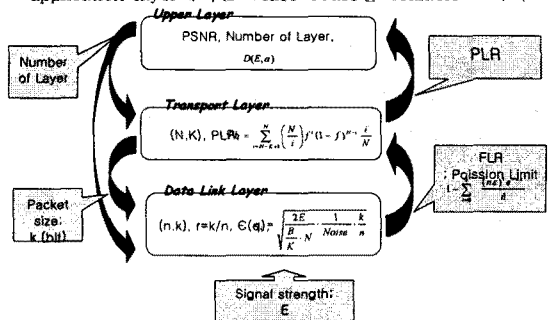
CDMA 채널별로 layer수를 다르게 하고 채널 Power allocation을 통한 video 전송을 제안한 Zhao의 의견과 구별된다.[1] 기존 연구 중에 세가지 계층을 모두 아우르는 Cross Layer에 관한 연구는 그리 많지가 않다.

본 논문에서는 SVC를 이용한 최적화된 Cross-Layer Video Streaming Algorithm을 제안하고자 한다.

2. 연구 배경

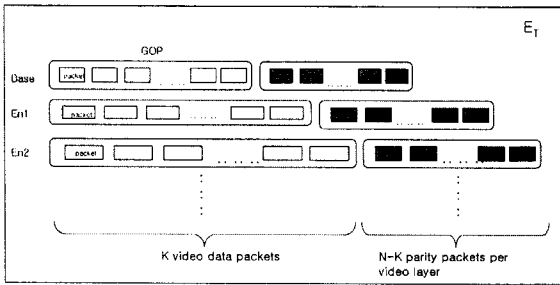
MPEG-4 표준규격에서 정한 layered coding 방식인 scalable video coding은 단순 coding 방식인 single layered 방식에 비교하여 error prone environment에서 현저한 성능 차이를 보여준다. CLO algorithm, Scalable video coding과 함께, 손실된 패킷을 어느 정도로 복원하는 FEC (Forward Error Correction) 방식을 같이 사용하여 비디오의 품질을 높일 수 있다. 본 논문에서 FEC coding은 (N, K)RS code를 사용하였다. K 개의 data packet과(N-K)개의 parity packet을 사용하여 N개의 packet에서 최대 (N-K)개의 Loss Packet까지 복원할 수 있다. Cross Layer는 layer 간의 상호 협력을 통한 기존 시스템의 성능 향상에 목적을 두고 있다.

application layer에서는 Video-codec을 scalable로 하여



<그림1> Cross Layer

QoS의 변화에 적용한다. QoS 변화는 사용자가 요구하는 방식과 네트워크의 환경변화에 적용하는 방식이 있다. 비디오 계층에서는 MPEG-4 Re-synchronization Marker를 사용하여 비디오 패킷간의 독립성을 보장한다. RTCP로 보고되는 네트워크 환경 변화에 따라 적절하게 비디오의 계층을 결정하는 것이 중요하다. transport layer에서는 RTP/RTCP를 이용하여 네트워크에서의 지연과 패킷손실율을 실시간으로 계산하여 미디어 QoS를 결정한다. FEC (forward error correction)로 Reed Solomon code를 이용하며, 패킷손실율에 따라 FEC 레벨을 결정한다. 본 논문에서는 Transport Layer 와 MAC Layer에서 FEC coding을 최적화하는 연구를 수행하였다. 또한, Burst한 loss시에는 retransmission이 적합하며, FEC와 retransmission 방식의 선택에 대한 알고리즘이 필요하다. Fading에 의한 네트워크 품질 저하에서는 FEC와 retransmission 등의 방식이 효과적이지만, congestion에 의한 네트워크 품질 저하시에는 단순히 비트 율을 낮춤으로서 대응한다. 이를 구별하기 위한 정보는 하위 Layer에서 제공되어야 한다. 본 논문에서 사용된 video source는 MPEG-4 Spatial SVC이며, 4개의 Layer로 구분하여 실험하였다. 그림 2에서 보는 바와 같이 하나의 GOP(Group of picture)는 20개의 Packet으로 이루어져있고 1초에 하나의 GOP를 각각의 Layer 별로 전송한다.



<그림2> Layer와 source/parity packet 간의 관계
각 Layer별 bitrate는 표 1과 같다.

Layer	bytes	bit k/sec
Base	15925	39.40
En1 Layer	34622	85.66
En2 Layer	88849	219.83
En3 Layer	176167	435.88

<표 1> Video Source bitrate/sec

Base-Layer에서 En3 Layer까지 모두 합하면 4QCIF급의 해상도(640×512)를 갖게 된다. 또한, 본 논문에서는 기존의 계층화 방식에서는 고급 계층 중에서 손실이 발생하면 그 계층 보다 상위에 있는 계층의 데이터를 사용하지 않고 있었지만, 손실이 일어난 계층의 상위계층을 활용하여 복원하는 것이 기존의 상위계층을 버리는 방법보다 화질 개선에 도움을 줄 수 있다는 연구를 바탕으로 실험하였다. SVC는 Layer가 도착 할수록 Distortion(= ΔD_i)이 작아지며 선택한 Layer에 대한 평균 Distortion은

$$\bar{D}(l) = \sum_{i \in l} P_i \cdot \Delta D_i \quad (1)$$

이다. l 는 선택한 레이어 set의 수이며 집합{B, 1, 2, 3}에서 B를 포함한 부분집합의 수와 같고 각각의 수는 1을 참, 0을

거짓으로 하는 2진수의 배열로서 Layer의 구성과 대응된다. l 는 선택된 l 에 따른 레이어 set이다. P_i 는 Transport Layer에서의 defFEC coding 후의 각 Layer set을 이루고 있는 GOP의 Loss Ratio로, 선택한 레이어마다 다르다. 각각의 Layer의 video source의 packet size는 다르며 선택된 Layer의 GOP당 Packet의 수는 R_i 로 표현된다. (본 논문에서는 Layer간 packet 수는 같게 하고 그 size를 다르게 하여 차별하였다)

Selected Layer	l	Layer set i	MSE
B	0	B, B1	343
B 1	1	B, B1	133
B 2	2	B, B2	208
B 1 2	3	B, B1 B2, B12	39
B 3	4	B, B3	285
B 1 3	5	B, B1 B3, B13	87
B 2 3	6	B, B2 B3, B23	165
B 1 2 3	7	B, B1, B2 B12, B13	7

<표2> Layer set

RS(N, K) code에서는 N-K개의 패킷까지 복원이 되므로

$$P_i = \sum_{R_i=N-K+1}^N \binom{N}{R_i} FLR^{R_i} (1-FLR)^{N-R_i} \frac{R_i}{N} \quad (2)$$

FLR은 MAC Layer에서 하나의 Packet을 이루고 있는 bit frame의 Loss ratio로서, 선택된 레이어의 하나의 패킷을 이루고 있는 bit가 k_i bit라고 할 때, bit 단위 FEC는 (n, k)라 할 수 있다. 이는 $n \rightarrow \infty, \epsilon_l \rightarrow 0$ 의 조건을 만족한다고 볼 수 있으므로 Poission Limit

$$f[i] = \frac{(n\epsilon_l)^i e^{-n\epsilon_l}}{i!} \quad \text{을 적용할 수 있다.}$$

(본 논문에서 $n=2000 \sim 22000$ [bit/packet]이다)

f 는 복원 확률이므로

$$FLR = 1 - \sum_{i=0}^{k_i} \frac{(n\epsilon_l)^i e^{-n\epsilon_l}}{i!} \quad (3)$$

가 된다. 본 논문에서는 PHY에서의 BER(bit error ratio)은 다음과 같이 나타냈으며, BPSK에 따르는 것을 가정하였다.

$$q = \sqrt{\frac{2E}{\frac{B}{N} \cdot N} \cdot \frac{1}{\text{Noise}} \cdot \frac{k}{n}} \quad (4)$$

본 논문에서는 SNR E를 분배 계수 α 에 따라 Layer 별로 다르게 분배하게 하여, \bar{D} 의 최소값을 Lagrange minimize에 의한 Optimal solution을 찾고 있다. A와 B는 error 확률이 0이나 1에 되지 않게 하기위한 SNR의 범위이며 a, b는 분배 계수 α 가 한 Layer라도 SNR이 0되지 않을 범위를 나타낸다.

$$\min \vec{D}(E) \text{ subject to } A \leq E \leq B, \quad (5)$$

$$\text{and } a \leq \alpha \leq b,$$

$$\left(\sum_{L=0}^3 e_L = E, (e_L - c_1) = \alpha(L - c_2) \right)$$

$$Y(E, \lambda) = \vec{D}(E) + \lambda E + c$$

$$= \sum_{i \in I} \left[R_i = \sum_{N=K+1}^N \binom{N}{R_i} FLR^{R_i} (1 - FLR)^{N - R_i} \frac{R_i}{N} \right] \cdot \Delta D$$

$$\lambda E + a \quad (6)$$

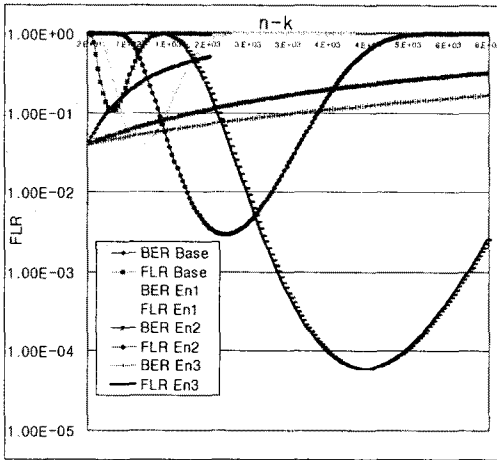
Lagrange minimize 수식은 Steepest Descent algorithm을 이용하여 최소값을 찾았다. 그러나 변수에 대해 편미분이 곤란하기 때문에 다음과 같이 편미분을 근사적으로 하는 수치 근사법을 이용하여 식을 전개하였다.

$$g = \frac{f(x+u, y) - f(x, y)}{u}, \quad (7)$$

$$h = \frac{f(x, y+u) - f(x, y)}{u}$$

3. 시뮬레이션 결과

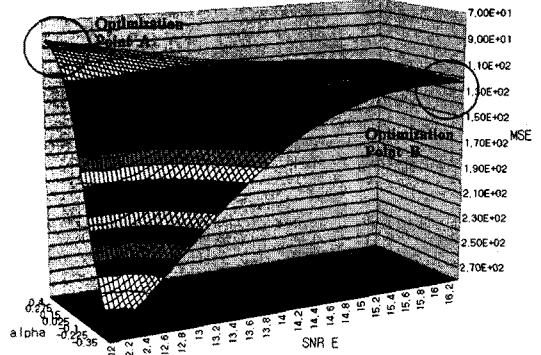
본 논문에서는 실제 실험되는 video source의 파라미터를 이용하여 보다 reality를 높였다. 사용된 콘텐츠 'foreman'은 Base에서 Enhance 3 layer(Layer set B123)일 때, 4CIF급의 해상도를 갖고 Base Layer(Layer set B)일 때, QCIF급의 해상도를 갖는다.



<그림3> 고정된 SNR에서 Layer별 최소 FLR과 parity bit

<그림3>은 Simulation 중간단계에서 각 레이어의 GOP마다 같은 SNR을 적용했을 때, MAC Layer에서의 최소 FLR이 구해지는 그림을 보여준다. 고정된 E에서는 parity bit를 너무 많이 붙이게 되면 오히려 bit당 E가 작아지 FLR이 늘어난다는 것과 상위 Layer에서 내려온 정보를 이용하여 각 레이어마다 최소 loss ratio를 위한 필요한 parity bit를 보여주고 있다. 본 논문에서는 SNR을 최소화 하면서, 그에 따라 trade off 관계에 있는 distortion을 최소화 하는 것에 목적이 있다. <그림5>는 레이어별 SNR의 분배와 SNR E에 따른 distortion을 보여준다. $\vec{D}(E)$ 가 최소가 될 가능성이 있는 부분을 Optimization Point A,B로 표시하였다. Steepest

Descent algorithm을 사용한 결과 최소치는 A point로 나타났다. 가중치 λ 을 얼마나 주는가에 따라 그래프의 모양이 변하지만 그에 따른 $\vec{D}(E)$ 을 찾을 수 있다. 이는 SNR을 최소화 하면서 최상의 Video Streaming을 할 수 있다는 것을 보여준다.



<그림5> Optimization Point

4. 결론

본 논문에서는 SCV와 FEC 기술을 통한 Cross Layer 기술로서 기존의 2계층 CL QoS가 아닌 3계층을 아우르는 Cross Layer Optimization을 통한 새로운 접근을 시도하였다. Video Source를 SVC하여 각 video Layer를 계층별로 GOP 단위로 전송하고, Network 레이아웃 정보를 교환 했을 때, 최소화 된 distortion을 구할 수 있었다. 또한, bitrate가 다른 Layer마다 각각 다른 distribution을 통하여 최상의 비디오 품질에 대한 전송 방법에 대해서도 살펴보았다.

이와 같은 Cross Layer Optimization 기술은 기존의 단일 계층 해법이 해결하지 못한 보다 넓은 영역의 문제와 최적화 문제를 다루는데 새로운 가능성을 제시한다.

참고 문헌

- [1] S. Zhao, Z. Xiong and X. Wang, "Joint Error Control and Power Allocation for Video Transmission over CDMA Networks with Multiuser Detection", *IEEE Trans. Circuits Syst. video Technol.* vol. 12, pp 425-437, JUNE 2002
- [2] R.Branden, D.Clark, and S.Shenker, "Integrated services in the Internet architecture: an overview." *RFC1633, IETF, June 1994*
- [3] R.Branden et al., "Resource reservation protocol(RSVP) version 1, function specification," *RFC2205, IETF, Sep. 1997.*
- [4] S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang, and W.Weiss, "An Architecture for Differentiated Services," *RFC2475, IETF, Dec. 1998.*