

적응 스트리밍을 위한 스케일러블 비디오 코딩 기법에 관한 연구

*박종순, 김동환, 장태규
중앙대학교 전자전기공학부
e-mail : tgchang@cau.ac.kr

Implementation of Scalable Video Coding method for adaptive streaming

*Jong-soon Park, Dong-hwan Kim, Tae-gyu Chang
School of Electronic and Electric engineering
Chung-ang University

Abstract

Adaptive streaming algorithm for the realtime video decoder that can be dynamically adaptable to the network channel status or terminal environment is described and performance analyzed. An LMS filter is used to predict the bitrate and predicted bitrate is used to optimize based prediction filter is used to adapt and optimize the quality of the decoded pictures. The quality of the output pictures are about 48dB.

I. 서론

네트워크 기반의 실시간 멀티미디어 스트리밍을 위해서는 네트워크의 변동을 보상하기 위한 기법이 필요하다. 이를 위하여 복수개의 서버를 이용하여 스트리밍을 수행하는 청크 알고리즘[1] 이나 CBR 을 보장하는 다양한 네트워크 QoS 방법들이 제시되어 왔다.

본 논문에서는 temporal, spatial, SNR 의 3가지 측면에서 확장성을 제공하는 Scalable Video Coding 표준을 위한 실시간 비트스트림 전송기를 적응 신호 처리 기법에 기반 하여 전송 대역폭의 변동 상황을 예측하

여 필요한 만큼의 확장 레이어 만을 전송하는 적응 스트리밍 알고리즘을 새로이 제시하고자 한다.

II. 관련 연구

2.1 Scalable Video Coding 표준

Scalable Video Coding(SVC) 는 H.264 Main Profile 을 인터넷이나 Wireless lan 과 같은 패킷 망에서 채널의 상황이나 디코더가 동작하는 단말의 성능 등에 적응적으로 동작할 수 있도록 확장하기 위하여 고안중인 표준이다. SVC 는 코덱 코드의 변화 없이 디코더 입력 비트스트림을 변화시키는 것 만으로 디코딩된 화면의 화질이나 프레임 레이트, 화면 크기의 제어가 가능한 3-way 확장성을 지원한다.[2] 아래 그림 1 에 SVC Decoder 의 블록 다이어그램을 표시하였다.

2.1.1 Temporal & Spatial Scalability

Haar wavelet MCTF 는 입력 화면을 시간축 과 공간축 상에서 low & highpass frame 과 프레임들의 모션 벡터로 변환하는 하는 기능을 수행한다[3]. 디코더에 있는 inverse MCTF 에서는 수신한 lowpass frame 과 highpass frame 을 합성하여 출력 화면을 복원하게 된

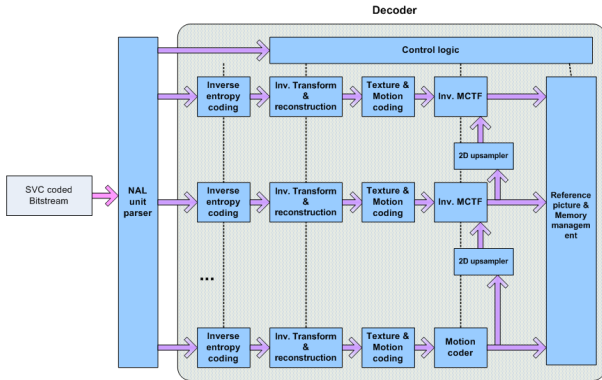


그림 1 여러 레벨의 확장성을 지원하는 Scalable Video Decoder 구조다. 이 과정에서 lowpass frame 에는 가장 작은 화면 크기에 최저의 화질정보만 저장되게 되며 highpass frame 에 의하여 최소한의 정보만 담기게 되며 highpass frame 이 덧붙여 질수록 시간축상에서 프레임의 숫자와 화면의 크기가 늘어나게 된다.

2.1.2 SNR scalability

확장 레이어의 DCT 계수값은 residual 신호의 공간 트랜스폼 계수이다. 이 계수는 양자화를 거친 후 CABAC 을 통하여 통계 압축을 수행하게 된다. SNR scalability 는 각 layer 마다 양자화 계수값을 달리하여 양자화를 수행하게 되므로 확장 레이어를 추가 할 때 마다 화질의 개선이 가능하다.

III. 적응 전송기의 구조 및 알고리즘

3.1 Traffic estimation

본 논문에서는 서버와 클라이언트가 연결되어 있는 망의 전송률 변화에 적응적으로 스트리밍을 위하여 서버에는 클라이언트의 수신 비트레이트를 예측하여 전송 가능한 만큼만 전송하도록 하였다. 이를 위하여 서버에는 Least Mean Square(LMS) 필터를 달았다. LMS 필터를 통하여 예측된 비트레이트에 따라 코딩된 비트스트림의 확장 레이어를 선택적으로 전송하도록 하였다.

3.1.1 LMS filter

LMS filter 는 입력값 $x(t)$ 에 따른 예측된 출력값 $y(t)$ 를 계산하는 filtering process 와 예측 오차를 최소화하기 위해 filter 의 계수값을 자동으로 조정하는 adaptive process 로 이루어져 있다[4]. 필터의 파라미터로는 필터의 탭수와 step size 파라미터 μ 가 있다. μ 는 필터의 adaptive process 의 진행속도를 결정하는 파라미터로서 이 값이 작으면 천천히 진행되지

만, gradient noise 가 적어지게 된다. 아래 그림에 필터의 탭수와 step size 파라미터에 따른 예측 오차에 대하여 나타내었다.

	1	4	8	16	32	64
0.1	10195	9130.2	10166	$3 \cdot 10^{26}$	$1 \cdot 10^{135}$	$1 \cdot 10^{134}$
0.05	10001	8734.3	8887.1	9924	$1 \cdot 10^{16}$	$7 \cdot 10^{71}$
0.01	9870.5	8531.2	8487.1	8732.2	9050.8	9520
0.005	9855.6	8511.2	8456.9	8674.5	8981.9	9192.7

표 3 탭수와 μ 에 따른 필터의 오차값

위의 표에서는 μ 값이 작아짐에 따라 예측오차가 줄어드는 것을 볼수 있다. 위의 표에서 가장 최적화된 값을 선택하여 이후의 simulation 과정을 진행하였다.

3.1.2 Traffic model

Realtime Transport Protocol(RTP) 와 같은 실시간 멀티미디어 데이터 전송 프로토콜은 대부분 connection-less(CL) 방식으로 데이터를 전송한다. 이는 CL 방식이 가지는 장점에 기인하고 있다. CL은 network 리소스에 대한 예약이 필요 없으므로 버스트한 데이터를 전송할 경우 connection oriented(CO) 방식에 비해 망에 대한 부하를 줄일 수 있다.[4]

MPEG 방식으로 코딩된 비트 스트림은 I 픽처가 B, P 픽처에 비해 크기가 크므로 아래의 그림과 같이 버스트한 트래픽을 유발하는 특성을 지니게 되며 전송도중 몇 개의 패킷이 유실되어도 전체 스트림의 디코딩에는 영향을 끼치지 않으므로 CL 방식이 CO 방식에 비해 비디오 비트스트림을 전송하기에 더욱 적합하다. 아래 그림에 MPEG 스트림의 프레임별 크기를 나타내었다. 네트워크 상에서 성능 검증을 위하여 트래픽을 모델링하였다. 네트워크 트래픽은 여러 가지의 변수들에 의해 변동을 보이게 되나 통계적인 성질로서 자기 유사성(Self similarity) 을 갖는다.[5] 이는 Fractional Brownian Motion 을 이용하여 모델링이 가능하며 다음의 식과 같이 정의된다.[4]

$$A_t = mt + \sqrt{am}Z_t, \quad t \in (-\infty, \infty) \quad (1)$$

위 식에서 m 은 평균 input rate, a 는 트래픽의 분산, Z_t 는 normalized fractional Brownian 트래픽이며 A_t 는 $[0, t)$ 구간의 트래픽의 양을 bit 로 표현한 값이다. 아래 그림 3에 발생시킨 트래픽과 이를 보상한 결과를 나타내었다.

3.2 적응 전송 알고리즘

적응 전송기는 LMS 필터를 이용한 트래픽 예측기가

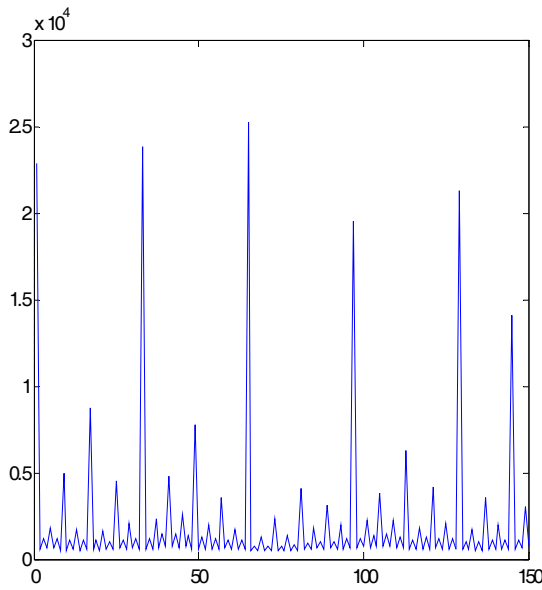


그림 2 SVC 비트스트림의 프레임별 크기(bytes) 예제

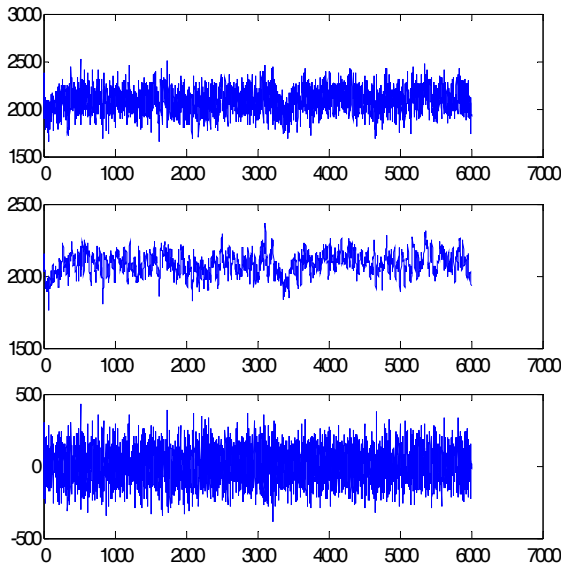


그림 3 fBm 을 이용하여 발생시킨 트래픽과 LMS filter 를 이용하여 보상한 결과

($m=2*10^3$, $a=1*10^7$, $tap=8$, $\mu=0.005$)
 있으므로 트래픽 예측기에 의해 예측된 트래픽의 양만큼 base 레이어와 확장 레이어를 클라이언트로 전송하면 된다. 그러나 이 방식의 문제점은 크기가 버스트하게 늘어나는 I 프레임의 경우 다른 프레임에 비해 확장 레이어를 많이 포기해야 하므로 화질의 열화가 심하게 일어나며 디코더에는 전송받은 비트스트림을

저장하기 위한 버퍼가 있으므로 이 버퍼에 의해 보상되는 네트워크 전송속도의 변동에 대한 고려가 빠져있다는 점이다. 이러한 문제점을 해결하기 위해서는 다음 번 시간 구간에 대한 전송률의 예측값 이외에도 버퍼에 있는 비트스트림의 양에 대한 고려가 필요하다. 따라서 n 번째 프레임에서 전송할 수 있는 확장 레이어의 크기는 n 번째 base 레이어의 크기와 클라이언트 비트스트림 버퍼의 양을 고려하여 결정하며 아래의 식과 같이 된다.

$$m[n] = (\hat{r}[n] + u[n-1]) - b[n] \quad (2)$$

위 식에서 $m[n]$ 은 전송 가능한 n 번째 확장 레이어의 크기이고, $\hat{r}[n]$ 은 예측된 전송률이며 $b[n]$ 은 n 번째 base layer 의 크기가 된다. $u[n-1]$ 은 n-1 프레임까지의 클라이언트의 비트스트림 버퍼에 남아 있는 bit 수이다. $u[n-1]$ 은 아래와 같이 정의된다.

$$u[n] = \begin{cases} 0 & (m[n-1] \leq enh[n-1]) \\ m[n-1] - enh[n-1] & (m[n-1] > enh[n-1]) \end{cases} \quad (3)$$

$enh[n]$ 은 n 번째 프레임의 확장 레이어 크기이다.

IV. 동작 성능 실험 결과

적용 전송 알고리즘의 동작 성능을 검증하기 위하여 네트워크 트래픽을 모델링하고, 이에 의해 발생하는 전송률 변동을 예측하기 위한 LMS 기반의 적응 필터를 설계하여 시뮬레이션을 수행하여 그 결과에 따라 적응 전송을 수행하였다. 실험 데이터는 CIF(352 x 288) 의 비트스트림을 사용하였다. 실험은 적응 전송된 결과 비트스트림과 원본 비트스트림사이의 화질 열화 정도를 Peak Signal to Noise Ratio(PSNR) 로 계산하였다. 아래 그림 4에 결과를 나타내었다. 그림 4의 아래 그림을 보면 버스트가 나타나는 주변에서 계단모양의 화질의 열화가 발생하는 것을 볼수 있다. 이는 버스트에 해당하는 프레임은 I 프레임과 base 레이어 프레임들이며 레퍼런스 프레임의 화질 열화에 따른 오차값이 그 주변 확장 레이어 프레임의 디코딩시 누적되기 때문이다.

V. 결론 및 향후 연구 방향

본 논문에서는 IP 망을 통하여 스트리밍된 멀티미디어 데이터의 재생 화질을 최적화 하기 위하여 위하여 네트워크의 전송 대역폭 변동을 LMS 기반의 적응 예측기로 예측함으로써 Scalable Video Codec 을 위한 적응 스트리밍 알고리즘을 새로이 제시하였다.

Ethernet Traffic (Extended Version), IEEE/ACM
Trans. on Networking, vol. 2, Feb. 1994.

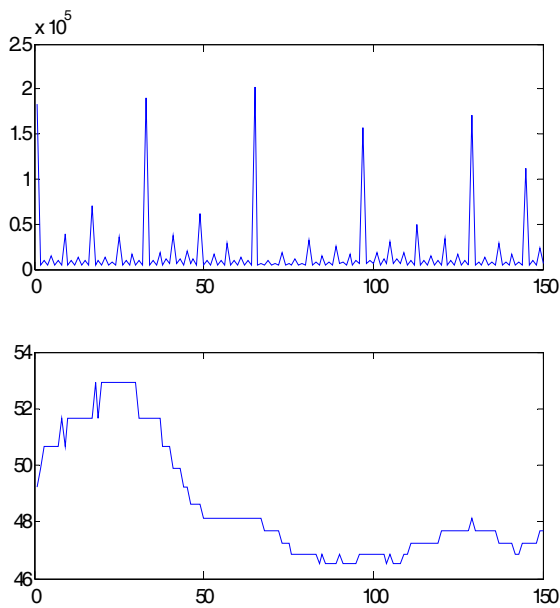


그림 4 테스트 비트스트림의 프레임별 크기와 디코딩 화면의 화질(PSNR)

적용 스트리밍 알고리즘의 성능을 검증하기 위하여 네트워크 트래픽을 모델링하고 LMS 기반의 8차 필터를 설계하였다.

향후 적용 스트리밍 알고리즘의 성능향상을 위해서는 레퍼런스 프레임의 오차가 주변 프레임들로 퍼져나가는 것을 막기 위해 베이스 프레임의 화질을 최적화 할 수 있는 알고리즘의 개발이 시급하다 하겠다.

참고문헌

- [1] D. Du, Reducing Bandwidth Requirement for Delivering Video over WAN with Proxy Server, IEEE Trans. on Multimedia, vol. 4, no. 4, Oct. 1999
- [2] ISO/IEC JTC 1/SC 29/WG 11, Scalable Video Model 3.0, October 2004.
- [3] M. Flierl, video coding with lifted wavelet transform and frame adaptive motion compensation, Proc. of VLBL, Sep. 2003.
- [4] S. HAYKIN, Adaptive Filter Theory, Prentice Hall(4th edition), 2002
- [4] I. Norros, On the Use of Fractional Brownian Motion in the Theory of Connectionless Networks, Selected Areas in Communications, IEEE Journal, 1995.
- [5] Will E. Leland, On the Self-Similarity Nature of