

오류 내성을 갖는 MPEG-4 부호화 기법

현 기 수, 문 지 용, 김 기 두, 강 동 옥
국민대학교 전자공학과
전화 : 02-910-4645 / 핸드폰 : 011-9349-7453

Error Resilient MPEG-4 Encoding Method

Ki-Soo Hyun, Ji-Yong Moon, Ki-Doo Kim, Dong-Wook Kang
Dept. of Electrical Engineering, Kookmin University
E-mail : whiz74@hotmail.com

Abstract

The main ideas of hybrid video coding methods are to reduce the spatial and temporal redundancy for efficient data compression. If compressed video stream is transmitted through the error-prone channel, bitstream can be critically damaged and the spatio-temporal error propagates through successive frames at the decoder because of drift noise in the references between encoder and decoder. In this paper, I propose the lagrangian multiplier selection method in the error-prone environment. Finally, it is shown that the performance comparisons of the R-D optimized mode decision are made against the conventional method and simulation results are given in the following.

I. 서론

기존의 동영상 코덱들을 비롯하여 MPEG-4 부호화기는 발생하는 데이터량을 줄이기 위해 공간적, 시간적, 통계적 중복성을 이용하고 이를 제거하기 위해 변환 부호화와 예측 부호화를 기반으로 하는 하이브리드 구조로 이루어져 있다. 이것은 ISO/IEC 뿐만 아니라 ITU-T 동영상 부호화 표준의 핵심이라 할 수 있다.

각 프레임을 일정한 크기의 블록으로 분할하고 이전

에 복원된 기준 영상과 현재 프레임의 시간적 상관성을 이용하여 각 블록에 대한 움직임 추정 및 보상을 행한 후 차분 신호를 얻게 되면 DCT(discrete cosine transform)를 이용하여 차분 신호를 주파수 영역으로 변환하게 된다. 그 후 양자화 과정을 통해 DCT 계수들을 양자화하고 움직임 추정을 통해 얻은 움직임 벡터를 각각 VLC(variable length coding)하여 비트스트림을 생성한다. 이러한 과정을 통해 생성된 비트스트림을 무선이동통신망을 통해 전송하게 되면 채널 오류 및 패킷 손실로 인해 복호 영상에 큰 영향을 주게 되므로 문제가 심각하다. 인터넷을 통해 비트스트림을 전송할 경우, 패킷 손실율이 약 20%에 달한다[1].

따라서, 채널 오류 및 패킷 손실에 의한 영향을 최소화하기 위한 오류 내성을 갖는 부호기가 중요한 문제로 대두되었다. 예측 부호화를 행하는 부호기를 통해 생성된 비트스트림이 손상을 받게 되면 그 오류의 영향이 현재의 프레임에만 국한되는 것이 아니라 그 후에 이어지는 프레임에도 계속해서 전파하게 된다[1, 6]. 이러한 전송 오류의 영향을 국부화시키기 위한 방법으로 인위적으로 인트라 블록을 삽입하는 방법과 동기 비트와 같은 추가적인 비트를 삽입하여 오류 전파를 차단하는 방법 등을 이용한다[2]. 전자의 경우, 채널 오류 및 패킷 손실에 의한 영향을 완화시킬 수 있는 방법으로 블록들에 대해 인트라 부호화를 하게 되면 이전의 블록과 현재의 블록간에는 종속성이 없어지게 되므로 오류의 전파의 영향을 막을 수 있다. 하지만, 인트라 부호화를 통해 부호화 했을 경우와의 비트율을 비교해보면 인트라 부호화의 경우 오류에 강인하다는 장점은 있는 반면, 발생하는 비트율은 상당히 높다는

단점이 있기 때문에 인트라 부호화를 많이 하게 되면 효율적인 압축이 목적인 동영상 코덱의 관점에서 볼 때 능률적이지 못하므로 효율적인 압축 및 오류에 대한 강인성을 위해 인트라 모드와 인터 모드를 적절한 방법을 통해 선택되어야 한다. 그러한 방법으로 사용 되는 것이 RD 최적화 기법이다[4]. 이 기법은 블록을 부호화 하면서 발생하는 비트율과 왜곡에 따른 RD 비용 함수를 이용하여 각 블록에 대한 최적화 부호 모드를 선택하게 된다. 동기 신호와 같은 추가적인 비트를 삽입하여 오류의 전파를 막는 후자의 경우, 수신된 비트스트림에 전송 오류가 발생하여 복호하는 과정에서 그 오류를 검출하게 되면 복호기는 다음의 동기 신호를 찾아낼 때까지 오류 은닉(error concealment) 기법을 이용하여 계속해서 영상을 재생해낸다. 일단 오류 은닉 과정을 거치게 되면 부호기내 가설 복호기에서 복원된 이전 프레임의 복호 영상과 복호기의 재생 영상과 차이로 인한 표류 잡음(drift noise)이 발생하여 다음 프레임으로 전파하게 된다[1,6,7].

본 논문에서는 채널의 전송 오류에 의해 복호기에서 오류 은닉이 이루어진다고 가정하고 그로 인해 발생하는 표류 잡음을 고려한 오류 내성 부호화 알고리즘으로서 MPEG-4 부호기가 MB 모드를 결정할 때 채널의 오류 특성에 의한 표류 잡음에 따라 적응적으로 선택되는 최적의 비용 함수를 제안하여 부호화 성능의 개선을 보인다.

II. 오류 내성을 갖는 MPEG-4 부호화기

전송 오류가 존재하는 채널을 통해 MPEG-4 부호기에 의해 생성된 비트스트림을 전송할 경우 채널 오류의 특성에 따라 비트스트림의 왜곡 정도는 다르지만 복호기는 이에 준하는 오류 은닉을 행하게 된다. 이러한 오류 은닉으로 인해 MPEG-4 부호기내의 가설 복호기를 통해 복원된 이전 프레임의 복호 영상과 복호기의 재생 영상간에 차이가 발생하게 되고 그것이 표류 잡음이 되어 전파하게 된다[1,6].

또한, 오류 내성 MPEG-4 부호기는 오류 은닉으로 인한 표류 잡음의 추정을 통해 인터 모드의 왜곡을 크게 하여 각 블록에 대한 모드 선택시 보다 많은 인트라 블록이 선택되도록 함으로써 오류 복원(error recovery)이 빨리 이루어지도록 한다. 그림 1은 MPEG-4 부-복호기에 적용된 알고리즘의 왜곡 모형을 보여준다.

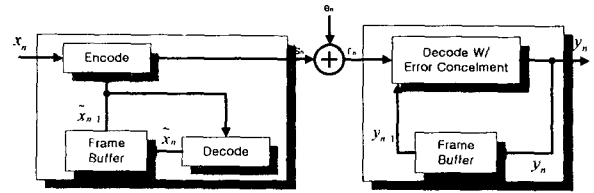


그림 1. MPEG-4 부호기와 복호기의 왜곡 모형

오류 내성 MPEG-4 부호기는 시간축에서의 예측 부호화와 공간에서의 변환 부호화를 행하는데 이에 따라 인트라 모드와 인터 모드로 나뉘어진다.

$$s_n = f(x_n, \hat{x}_{n-1}) = \begin{cases} f_r(x_n, \hat{x}_{n-1}) & \text{INTER} \\ f_a(x_n) & \text{INTRA} \end{cases} \quad (1)$$

여기서, x_n 은 입력 영상, \hat{x}_{n-1} 은 부호기내의 가설 복호기를 통해 복원된 영상, $f(x_n, \hat{x}_{n-1})$ 는 인트라 혹은 인터 모드를 행하는 함수, s_n 은 부호기에 의해 생성된 비트스트림이다.

비트스트림에 전송 오류가 검출되면 복호기에서는 시간축에서의 오류 은닉(temporal error concealment), 움직임 보상을 이용한 오류 은닉(motion compensated error concealment)을 행하며 전송 오류가 없을 경우는 정상적으로 복호하며 다음과 같이 표현된다.

$$y_n = \begin{cases} y_{n-1} & \text{if } TEC \\ y_{n-1}(mv) & \text{if } MCEC \\ f^{-1}(r_n, y_{n-1}) & \text{otherwise} \end{cases} \quad (2)$$

각각의 오류 은닉 기법에 따른 표류 잡음의 전파는 다음과 같다.

$$d_n = y_n - \hat{x}_n = \begin{cases} y_{n-1} - \hat{x}_n & \text{if } TEC \\ y_{n-1}(mv) - \hat{x}_n & \text{if } MCEC \\ f^{-1}(s_n, y_{n-1}) - \hat{x}_n & \text{otherwise} \end{cases} \quad (3)$$

여기서, TEC가 일어날 확률을 P_{TEC} , MCEC가 일어날 확률을 P_{MCEC} 라 하면 양자화 잡음 및 표류 잡음에 의한 총 왜곡은 인트라 모드와 인터 모드에 따라 다음과 같이 표현된다.

$$D_r = P_{TEC} \cdot |y_{n-1} - x_n|^2 + P_{MCEC} \cdot |y_{n-1}(mv) - x_n|^2 + (1 - P_{TEC} - P_{MCEC}) \cdot |f_r^{-1}(s_n, y_{n-1}) - x_n|^2 \quad (4)$$

$$D_a = P_{TEC} \cdot |y_{n-1} - x_n|^2 + P_{MCEC} \cdot |y_{n-1} - x_n|^2 + (1 - P_{TEC} - P_{MCEC}) \cdot |f_r^{-1}(s_n) - x_n|^2 \quad (5)$$

여기서, D_r 와 D_a 는 표류 잡음을 고려한 인터 모드와 인트라 모드의 총 왜곡을 각각 의미한다. 또한, 양자화 잡음과 표류 잡음이 서로 독립적이라 가정하면 식 (5), (6)은 다음과 같이 표현된다.

$$D_{total} = D_{drift} + D_{quant} = \|d_n\|^2 + \|q_n\|^2 \quad (6)$$

여기서, d_n 은 표류 잡음, q_n 은 양자화 잡음이다 [6,7].

III. 표류 잡음을 고려한 MPEG-4 부호기에서의 lagrangian multiplier 선택 방법

효율적인 압축을 위해 전송 오류가 없다고 가정하면 각 MB의 블록 모드를 인터 모드로 부호화하는 것이 가장 좋다. 하지만, 전송 오류가 일어날 확률이 높은 채널을 통해 비트스트림을 전송하게 되면 패킷 손실 및 전송 오류로 인한 비트스트림의 변형이 일어나게 되므로 부호기에서는 오류에 대한 강인성을 고려하여 인트라 모드와 인터 모드를 적절하게 선택하여야 한다. 이러한 부호 모드 선택은 일반적으로 RD 비용 함수를 이용하여 최적화된다.

$$J(R) = D(R) + \lambda R \quad (7)$$

여기서, λ 는 lagrange 변수이며 다음과 같이 나타낸다[4].

$$\lambda_0 = 0.85 \cdot Q^2 \quad (8)$$

기존의 λ_0 를 유도할 때 D는 모든 구간에서 미분 가능하며 양자화 잡음은 균일한 확률 분포를 가지고 있다고 가정을 하였으며 식 (9)와 (10)를 이용하여 얻어졌다.

$$R(D) = \alpha \ln\left(\frac{\sigma^2}{D}\right) \quad (9)$$

$$D_{quant} = \frac{Q^2}{3} \quad (10)$$

여기서, α 는 입력 영상의 pdf에 따른 값이고, σ^2 는 입력 영상의 분산이다[4,7].

하지만, 채널의 전송 오류가 발생할 경우 총 왜곡은 식 (6)과 같이 양자화 잡음뿐만 아니라 표류 잡음에 의한 왜곡도 고려해야 한다. 양자화 잡음과 표류 잡음이 서로 독립적이라 가정하고 식 (6), (9), (10)을 Q에 관해 각각 미분하면 다음과 같은 결과를 얻을 수 있다.

$$dR = \frac{-2\alpha Q}{Q^2 + 3D_{drift}} dQ \quad (12)$$

$$dD = \frac{2Q}{3} dQ \quad (13)$$

그리고, 식 (12)와 (13)을 정리하면 다음과 같이 나타낼 수 있다.

$$\frac{dD}{dR} = -\frac{1}{3\alpha}(Q^2 + 3D_{drift}) \quad (14)$$

그리고, λ 는 식 (7)을 R에 관해 미분하면

$$\lambda = -\frac{dD}{dR} \quad (15)$$

이므로, 식 (14)와 (15)를 정리하면

$$\lambda = \frac{1}{3\alpha} Q^2 + \frac{1}{\alpha} D_{drift} \quad (16)$$

또한, $\lambda_0 = \frac{1}{3\alpha} Q^2$ 라 하면 α 는 대략 0.39가 된다.

최종적으로 표류 잡음을 고려한 lagrange multiplier는 다음과 같이 표현할 수 있다.

$$\lambda \approx \lambda_0 + 2.55 \cdot D_{drift} \quad (16)$$

IV. 제안하는 lagrangian multiplier를 이용한 부호화 성능 분석

본 논문에서 제안한 λ 와 패킷 손실 확률이 $3e-02$ 인

인터넷에서의 패킷 손실 패턴을 이용하여 오류 내성 MPEG-4 부호기의 성능을 분석하였다. 모의 실험에서 이용된 qcif 영상은 foreman이다. 기존의 lagrange multiplier를 이용했을 경우와의 성능 비교를 위해 제안한 λ 를 이용한 오류 내성 MPEG-4 부호기를 통해 부호화된 비트스트림을 채널 시뮬레이터를 이용하여 오류 패턴을 삽입한 후 복호기를 통해 복원된 영상의 SNR과 부호기에서의 비트율을 비교 분석하였다.

그림 2~5는 Q가 각각 12, 16, 20, 24, 28이고 BER이 각각 0.0001, 0.0003, 0.0005, 0.0007일 경우 기존의 방법과 제안한 λ 를 이용한 오류 내성 MPEG-4 부호기의 성능을 보여준다.

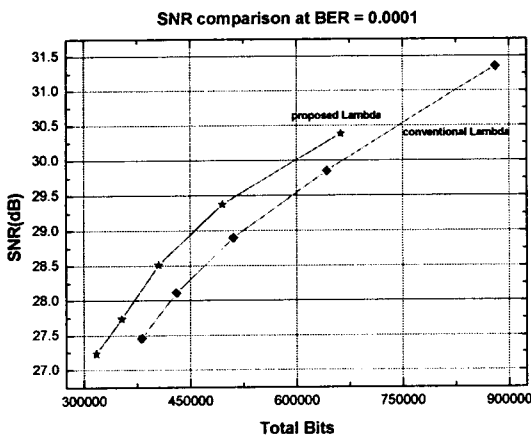


그림 2. 제안한 λ 를 이용한 BER = 0.0001에서의 SNR 비교

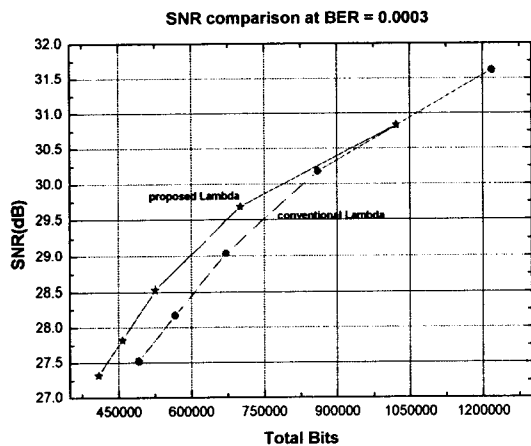


그림 3. 제안한 λ 를 이용한 BER = 0.0003에서의 SNR 비교

기존의 방법에 비해 제안한 방법을 이용할 경우 동일한 비트율에서 약 0.1~1dB 정도의 성능 개선을 볼 수 있으며, 동일한 SNR에서 볼 때 BER이 0.0007인 경우 많은 비트율 저감 효과를 얻을 수 있었다.

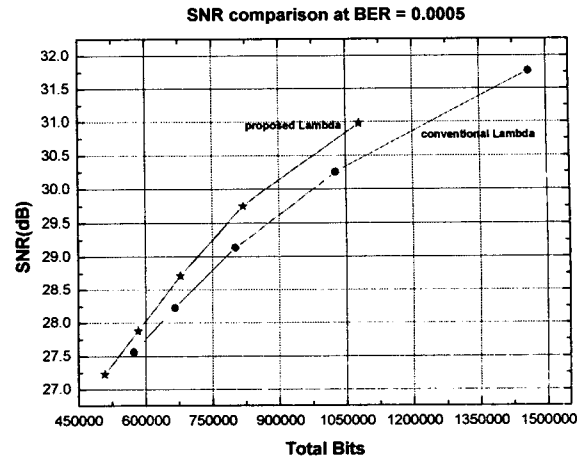


그림 4. 제안한 λ 를 이용한 BER = 0.0005에서의 SNR 비교

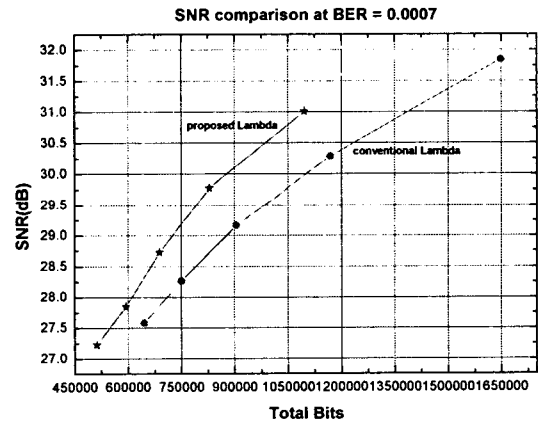


그림 5. 제안한 λ 를 이용한 BER = 0.0007에서의 SNR 비교

IV. 실험결과 및 고찰

본 논문에서는 채널에 의한 전송 오류로 인해 발생하는 표류 잡음을 고려한 lagrange multiplier를 제안하여 성능의 개선을 보였다. 인터넷 패킷 오류가 일어날 확률이 $3e-2$ 일 경우 약 0.1~1dB의 성능 개선을 확인

할 수 있었으며 많은 비트율 저감 효과도 있음을 알 수 있었다. 앞으로 여러 가지 qcif 영상과 wcdma에서의 오류 패턴들을 이용한 추가적인 실험을 통해 제안한 λ 의 우수성을 보일 것이다.

Reference

[1], Rui Zhang, Shankar L. Regunathan and Kenneth Rose, "Video Coding with Optimal Inter/Intra-Mode Switching for Packet Loss Resilience", IEEE Journal on Selected Areas In Communications, VOL 18, NO.6, JUNE 2000.

[2] Yao Wang, Stephan Wenger, Jiangtao Wen and Aggelos K. Katsaggelos, "Error resilient video coding techniques", IEEE Signal Processing Magazine, pp.61-82, July 2000.

[3] ISO/IEC JTC1/SC29/WG/11 N2459, "MPEG-4 Overview", Oct. 1998.

[4] Gary J. Sullivan and Thomas Wiegand, "Rate-Distortion Optimization for Video Compression", IEEE Signal Signal Processing Magazine, pp 74~90, Nov. 1998.

[5] Thomas Stockhammer et al., "Optimized transmission of H.26L/JVT coded video Over packet-lossy Networks".

[6] Chul-Woo Kim et al., "High complexity mode decision for error prone channel", ITU-SG16/Q6 Contribution: VCEG-043, Nov. 2001.

[7] Chul-Woo Kim et al., "High complexity mode decision for error prone channel", ITU-T SG16/Q6 Contribution: JVT-C101, May 2002.

[8] T. Wiegand and B. Girod, "Lagrange Multiplier Selection In Hybrid Video Coder Control", ICIP'01.