

H.264/AVC를 위한 에러 적응형 예비 슬라이스

*이진호 **최진수 **최해철

*과학기술연합대학원대학교, **한국전자통신연구원

jinosoul@etri.re.kr

Loss-Aware Rate-Distortion Optimization for Redundant Picture in H.264/AVC

*Lee, Jinho **Choi, Jin Soo **Choi, Haechul

*University of Science and Technology, **Electronics and Telecommunications Research Institute

요약

예비 슬라이스(Redundant picture)는 에러가 발생하기 쉬운 환경에서 비디오가 전송될 때 에러를 보상하기 위한 H.264/AVC 부호화 표준의 도구들 중에 하나이다. 본 논문에서는 예비 슬라이스의 삽입으로 인해 발생하는 오버헤드 비트와 손실률을 고려한 LA-RDO(Loss-aware rate-distortion optimization) 모델을 제시하며 제안하는 모델에 기초해서 예비 슬라이스의 삽입 여부를 결정한다. 실험결과는 제안하는 방법이 참조 소프트웨어(JM : Joint Model)에 있는 HRP(Hierarchical redundant picture)와 비교했을 때 평균 0.5dB의 PSNR이 증가함을 보였다. 특별히 낮은 비트율에서 상대적으로 높은 성능을 보이므로 비디오 화상 회의와 같이 낮은 비트의 영상을 요구하는 무선 전송 환경에 적합할 것으로 보인다.

1. 서론

무선과 같은 에러 발생이 가능한 네트워크 환경에서의 실시간 비디오 전송은 압축효율과 네트워크 적응력을 고려해야 하는 도전적인 과제이다. 이러한 요구들을 만족하기 위해 ITU-T VCEG과 ISO/IEC MPEG의 JVT는 개선된 부호화 효율과 네트워크 적응력에 초점을 맞춘 H.264/AVC[1] 비디오 표준을 개발해왔다.

에러 발생 가능한 환경에서의 비디오 전송시에는 패킷 손실이나 인코더와 디코더간의 참조 프레임 불일치의 문제들이 발생할 수 있다. 게다가 이러한 문제들은 에러 전파를 유발하고 전파된 에러들은 인트라 픽처와 같이 영상의 부호화가 새롭게 시작되는 지점(refresh)을 만나기전까지 축적되어진다.

H.264/AVC 부호화 표준의 베이스라인 프로파일에서는 이러한 현상들을 해결하기 위한 에러내성용 기술요소들[2] 중의 하나로 예비 슬라이스를 지원한다. 이 방법은 프라이머리 픽처(primary picture)라고 불리는 부호화 된 영상과 예비 슬라이스를 같이 전송하여 에러가 발생했을 때 손실된 프라이머리 픽처 자리를 예비 슬라이스로 대체함으로써 에러를 보상한다. 간단한 예비 슬라이스 구현 방법은 원영상의 복제본을 사용하는 것인데, 추가되는 예비 슬라이스의 개수에 비례하여 비트수가 증가하기 때문에 부호화 효율은 감소하게 된다. 따라서 많은 양의 예비 슬라이스를 삽입한다면 오버헤드 비트율이 급격하게 증가하게 되므로 제한된 처리량을 가지는 네트워크 환경에서는 유용하지 않을 것이다.

본 논문에서는 추가되는 비트의 양을 최대한 줄이면서 적응적이고 효과적인 예비 슬라이스 사용 방법을 소개하고자 한다. 이 전의 연구[3]에서는 예비 슬라이스의 위치를 계층적으로 할당하는 HRP가 제안됐고 참조 소프트웨어에 추가되었다. HRP는 예비 슬라이스의 위치를 GOP(Group of picture)안에서 계층적인 위치에 고정함으로써 축적되어진 에러를 주기적으로 보상하고 은폐(concealment)하는 방법이다. 이 방법에서는 예비 슬라이스의 개수를 각 GOP 단위 마다 한 개, 두 개 또는 세 개로 초기 설정할 수 있다.

그러나 HRP는 물체의 움직임이나 장면의 변화 같은 비디오 콘텐츠의 특성들을 고려하지 않은 상태에서 예비 슬라이스의 삽입 여부가 결정되기 때문에 이러한 영상의 특성들을 효과적으로 반영한 모델을 사용하여 예비 슬라이스를 삽입 한다면 그로 인해 발생하는 비트의 양을 줄이는데 도움을 줄 수 있을 것이다.

[4]에서의 율-왜곡(Rate-Distortion) 모델은 에러 발생 가능한 환경에서 알맞은 매크로블록(MB : Macroblock) 부호화 옵션을 선택하는데 효과적인 방법으로 제안되었다. 이 모델에서의 End-to-end 예측[4]은 손실되는 매크로블록에 의해 발생하는 잠재적으로 축적된 에러 왜곡을 예측한다. 또한 이 모델은 매크로블록의 특성들을 효과적으로 파악하여 전송효율과 계산 복잡도면에서 두드러진 성능 향상을 보였다. 마찬가지로 이 모델의 End-to-end 왜곡 예측을 픽처 기반에 맞게 개정하면 예비 슬라이스를 알맞은 위치에 할당하는데 도와줄 수 있을 것이다. 결론적으로 본 논문에서는 패킷 손실 환경에서 영상의 특성

들을 분석하는 개정된 LA-RDO 모델을 제안하고자 한다. 그리고 이 분석은 예비 슬라이스를 효과적으로 할당할 수 있게 도울 것이다.

논문의 2절에서는 적응적인 예비 슬라이스 할당을 위해 제안하는 올-왜곡 모델과 개정된 End-to-end 왜곡 예측을 상세하게 기술한다. 그리고 3절에서는 제안하는 방법과 참조 소프트웨어에 포함되어 있는 HRP, LA-RDO를 각각 실험하여 다양한 관점에서 성능을 비교하며 마지막 4절에서 결론으로 본 논문을 맺는다.

2. 제안하는 올-왜곡 모델

가. 올-왜곡 최적화 기법

올-왜곡 최적화(RDO : Rate-Distortion Optimization) 기법은 매크로 블록 모드 결정, 비율 조절(Rate control), 에러 적응형 올-왜곡 최적화(LA-RDO)[5] 기법 등 H.264/AVC의 많은 부분에서 성능향상을 위해 효과적으로 사용되어왔다. 에러가 없는 환경에서 한 프레임의 올-왜곡 비용값은 다음과 같이 나타낸다.

$$J(n) = D(n) + \lambda_o R(n), \quad (1)$$

식 (1)에서 $D(n)$ 과 $R(n)$ 은 각각 n 번째 프레임에서의 왜곡과 올을 나타낸다. 그리고 λ_o 는 라그랑제 승수이며 양자화 파라미터(QP : Quantization parameter) $Q[6]$ 에 의해 다음과 같이 표현되어진다.

$$\lambda_o = 0.85 \cdot 2^{Q/3}, \text{ for } H.264. \quad (2)$$

그러나 패킷 손실이 가능한 환경에서는 에러 전파에 의한 채널 왜곡이 발생되며 이러한 현상은 위에서 설명한 올-왜곡 모델의 사용을 부적절하게 만든다.

나. 왜곡 예측

에러가 발생하기 쉬운 환경에서의 왜곡을 효과적으로 예측하기 위해서는 현재 프레임이 손실되었을 때와 손실 되지 않았을 때의 두 가지 경우를 고려해야 한다. 첫 번째의 경우로 현재 프레임이 손실 되었다면, 손실 된 에러를 은폐하고 난 후에 발생하는 error-concealed 왜곡이 있으며 이 왜곡은 인트라 프레임 같은 부호화가 새로 시작되는 지점을 만날 때까지 에러를 전파 할 것이다. 두 번째 경우로 현재 프레임이 손실되지 않았다면, 에러가 없는 환경에서 발생하는 왜곡과 같은 error-free 왜곡이 발생되고 또한 이전의 프레임들로부터 현재 프레임까지 축적되어져 온 error-propagated 왜곡이 발생할 것이다. 종합적으로 n 번째 프레임에서 발생 가능한 모든 왜곡은 다음과 같다.

$$D(n) = (1-p)(D_{ef}(n) + D_{ep}(n)) + pD_{ec}(n), \quad (3)$$

위의 식에서 p 는 패킷 손실률(PLR : Packet Loss Rate)을 나타낸다. $D_{ef}(n)$ (그림 1에서 ①)과 $D_{ep}(n)$ (그림 1에서 ②)은 각

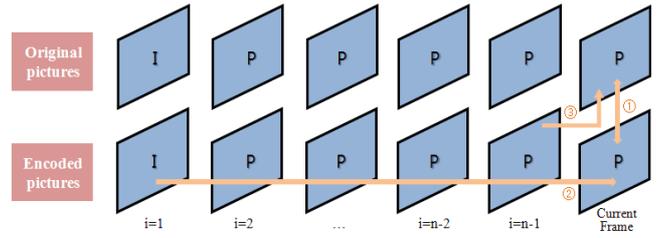


그림 1. 비디오 시퀀스의 구조

각 error-free 왜곡과 잠재되어 있는 error-propagated 왜곡을 나타내며 $D_{ec}(n)$ (그림 1에서 ③)은 error-concealed 왜곡을 나타낸다. 여기서 $D_{ep}(n)$ 은 직접 계산되어지지 않고 이전 프레임에서의 왜곡인 $D_{ep}(n-1)$ 와 $D_{ec}(n-1)$ 에 의해 아래의 식과 같이 유도되어진다.

$$D_{ep}(n) = (1-p)D_{ep}(n-1) + pD_{ec}(n-1), \quad (4)$$

식 (4)에서 $D_{ep}(n-1)$ 과 $D_{ec}(n-1)$ 은 $(n-1)$ 번째 프레임에서의 error-propagated 왜곡과 error-concealed 왜곡을 나타내며 $D_{ep}(n)$ 는 다시 다음 프레임으로 전파된다.

위에서 말한 것과 같이 비디오 시퀀스에서 발생한 에러는 다음 프레임으로 전파된다. 따라서 하나의 프레임이 아닌 모든 프레임의 패킷 손실률을 고려하여 식 (3)을 확장해야 된다. 그림 1은 비디오 시퀀스의 구조이다. 시퀀스의 시작에 해당하는 인트라 프레임이 에러 없이 전송되었다는 가정 하에 현재 프레임에서의 총 왜곡을 공식화 할 수 있다.

$$D_t(n) = (1-p)\{(1-p)^{n-1}D_{ef}(n) + D_{ep}(n)\} + pD_{ec}(n), \quad (5)$$

여기서 $D_t(n)$ 은 n 번째 프레임에서의 총 왜곡이다. $i=1$ 부터 현재 프레임까지의 모든 프레임이 에러 없이 전달되어졌다면 $D_{ef}(n)$ 을 구할 수 있다. 그리고 $D_{ep}(n)$ 은 $i=1$ 부터 현재 프레임까지 식 (4)에 의해서 축적되어진다. 주목할 점은 $D_{ep}(n)$ 을 구하는 식에 이미 이전에 전파되어진 왜곡 $D_{ep}(n-1)$ 과 $D_{ec}(n-1)$ 을 포함하기 때문에 $D_{ec}(n)$ 을 구하는데 있어서 이전 프레임들의 왜곡이나 패킷 손실률을 고려하지 않는다.

이제 마지막으로 예비 슬라이스를 고려한다. 만약 프레임 픽처와 예비 슬라이스의 패킷 손실률이 같다고 가정하면 식 (5)에서 예비 슬라이스의 패킷 손실률만 삽입함으로써 간단하게 표현 할 수 있다.

$$D_t(n) = (1-p \cdot p^{\epsilon_c}) \left\{ \left(\prod_{i=1}^{n-1} (1-p \cdot p^{\epsilon_i}) \right) D_{ef}(n) + D_{ep}(n) \right\} + p \cdot p^{\epsilon_c} D_{ec}(n), \quad (6)$$

ϵ_c 와 ϵ_i 는 현재 프레임(current frame)과 i 번째 프레임에서의 예비 슬라이스 유무를 나타낸다(예비 슬라이스의 삽입/미삽입 : $\epsilon = 1/0$). 예비 슬라이스는 프레임 픽처를 복사하는 간단한 방법으로 부호화 된다고 가정하였기 때문에 예비 슬라이스의 왜곡은 프레임 픽처의 왜곡과 같다고 본다.

다. 예비 슬라이스의 삽입 결정

이번 섹션에서는 제안하는 올-왜곡 모델을 어떠한 기준에 의해서 사용함으로써 예비 슬라이스의 삽입 여부를 결정할지를 설명하고자 한다. 우선 올-왜곡 비용을 구하기 위해서 올과 라그랑제 승수를 구한다. n 번째 프레임에서의 예비 슬라이스를 고려한 올은 다음과 같다.

$$R_t(n) = (1-p \cdot p^{\epsilon_c}) \left(\prod_{i=1}^{n-1} (1-p \cdot p^{\epsilon_c}) \right) R(n), \quad (7)$$

$R_t(n)$ 은 n 번째 프레임에서의 총 올을 나타내며 식 (6)에서와 같이 n 번째 슬라이스와 현재 슬라이스의 패킷 손실률을 고려하여 구한다. 그리고 패킷 손실 전송환경에서의 라그랑제 승수는 [4]에 의해서 구할 수 있다.

$$\lambda = (1-p \cdot p^{\epsilon_c}) \lambda_0, \quad (8)$$

여기서 λ_0 는 식 (2)에 나타나 있는 에러가 발생되지 않는 환경에서의 라그랑제 승수를 나타내며 최종적으로 식 (6)과 식 (7), 식 (8)을 식 (1)에 대입함으로써 에러 적응형 올-왜곡 최적화 비용을 구할 수 있다.

$$\begin{aligned} J(n) &= D_t(n) + \lambda R_t(n) \\ &= (1-p \cdot p^{\epsilon_c}) \left\{ \left(\prod_{i=1}^{n-1} (1-p \cdot p^{\epsilon_c}) \right) D_{ef}(n) + D_{ep}(n) \right\} \\ &\quad + p \cdot p^{\epsilon_c} D_{ec}(n) \\ &\quad + (1-p \cdot p^{\epsilon_c})^2 \cdot \lambda_0 \cdot \left(\prod_{i=1}^{n-1} (1-p \cdot p^{\epsilon_c}) \right) R(n), \quad (9) \end{aligned}$$

현재 프레임에 해당하는 예비 슬라이스를 부호화 하는 여부를 결정하기 위해 $J_1(n)$ 과 $J_2(n)$ 을 매 프레임 비교한다. 이 두 비용들은 각각 현재 프레임의 예비 슬라이스가 부호화 되었을 경우($\epsilon_c = 1$)와 부호화 되지 않았을 경우($\epsilon_c = 0$)의 올-왜곡 비용을 나타낸다. 올-왜곡 이론에서는 작은 비용을 가지는 방법이 최적의 방법이므로 아래의 식이 만족할 때 예비 슬라이스를 부호화 한다.

$$J_1(n) < J_2(n). \quad (10)$$

3. 실험 결과

제안하는 모델의 타당성을 보이기 위해서 ITU-T에서 권고하는 H.264/AVC 참조소프트웨어 JM 13.0에 적용하여 실험하였으며 이전의 두 가지 방법, 즉 LA-RDO(예비 슬라이스를 포함하지 않는 방법)와 HRP를 비교하였다.

예비 슬라이스는 H.264/AVC 베이스라인 프로파일에 포함되어 있고 모바일 멀티미디어나 비디오 화상회의에 사용된다. 따라서 모바일 서비스의 특성들을 고려하여 아래와 같이 테스트 조건을 설정하였다.

- Test Sequence : Carphone, News (qcif).
- FrameRate : 10fps.

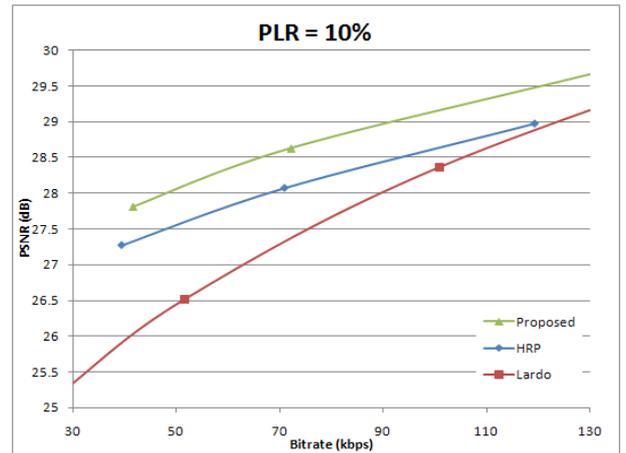


그림 2. 'Carphone' 영상의 비트율에 따른 PSNR 그래프.

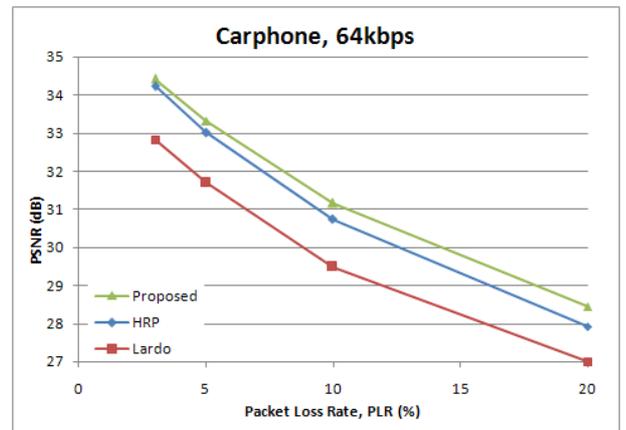


그림 3. 'Carphone' 영상의 패킷 손실률에 따른 PSNR 그래프.

- IntraPeriod : 60frames.
- Quantization Parameter : 20~32
- For HRP, NumRedundantHierarchy = 0.

인터넷 형태의 에러 패킷[7]과 같이 임의적으로 발생하는 에러와 그로 인해 악화되는 비트 스트림의 환경에서 비트율과 PSNRs의 관계를 효과적으로 실험하기 위해 비디오 부호화 표준 개발에 공통으로 사용되는 각 테스트 영상을 반복적으로 붙여서 5000 프레임의 영상을 만들었다.

그림 2는 패킷 손실률이 10%일 때의 비트율에 따른 PSNRs 그래프를 나타내는 것으로 모든 범위에서 제안한 모델의 성능이 다른 방법들보다 상대적으로 뛰어난 것을 볼 수 있다. 특별히 비트율이 낮을 때 제안한 모델과 LA-RDO와의 격차가 증가함을 보였다. 또한 패킷 손실률의 관점에서의 실험 결과로, 그림 3은 특정 비트율(64kbps)에서의 패킷 손실률 3%, 5%, 10% 그리고 20%에 대한 상기 기술한 방법들의 평균 PSNRs 그래프를 나타낸다. 이 그래프에서도 그림 2에서와 같이 평균 PSNRs가 상대적으로 뛰어난 양상을 보였다.

제안한 방법으로 인해 에러 보상측면에서의 향상을 나타내는 그래프를 그림 4에서 보여준다. 이 그래프는 각 프레임에서의 제안한 모델과 HRP의 객관적인 PSNRs 성능을 나타낸다. 예를 들어, 90 번째 프레임에서 HRP의 PSNRs는 에러의 영향

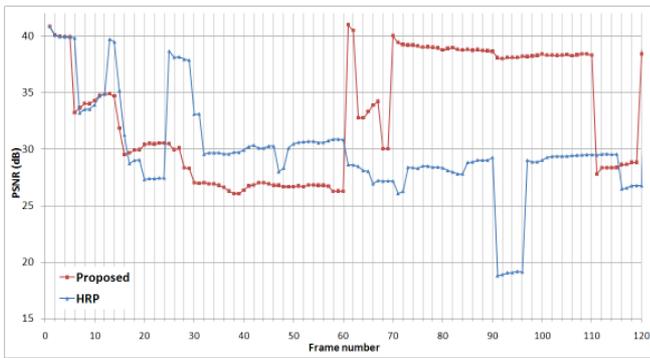


그림 4. 'News' 영상의 패킷 손실률이 10%일 경우 프레임에 따른 휘도 PSNR 그래프.



(a) (b) (c)

그림 5. 제안한 모델의 성능.

(a) 복원된 영상 (b) HRP에 의한 영상 (c) 제안된 영상

으로 급격하게 감소하게 되고 따라오는 프레임까지 에러가 전파됨을 보인다. 반면에 제안한 방법은 에러를 효과적으로 보상하며 따라오는 프레임들의 화질을 보호해 준다.

마지막으로 그림 5는 그림 4에서 70 번째 프레임을 나타내는데 (b)는 에러의 영향으로 화질이 악화된 반면에 제안한 방법은 주관적인 화질에서도 우수한 성능을 보였다.

4. 결론

본 논문에서는 패킷 손실 환경에서의 적응적인 예비 슬라이스 할당을 위한 효과적인 율-왜곡 모델을 제안했다. 이 제안된 방법은 잠재되어 있는 error-propagated 왜곡과 입력 비디오 시퀀스 특성들 사이의 고유한 상관관계를 나타내준다. 더 나아가 이 방법은 모바일 폰과 같은 낮은 비트율의 전송 서비스에 효과적으로 적용될 수 있음을 보였다.

5. 감사의 글

본 연구는 지식경제부 및 정보통신연구진흥원의 IT원천기술개발사업의 일환으로 수행하였음. [2008-F-011-01, 차세대 DTV 핵심기술 개발]

This work was supported by MKE/IITA [2008-F-011-01, Development of Next-Generation DTV Core Technology]

6. 참고문헌

- [1] ITU-T and ISO/IEC JTC 1, Advanced video coding for generic audiovisual services, ITU-T Recommendation H.264 and ISO/IEC 14496-10 (MPEG4-AVC).
- [2] T. Stockhammer, M.M. Hannuksela, and T. Wiegand, "H.264/AVC in Wireless Environments," *IEEE Trans. on CSVT*, vol. 13, no. 7, Jul. 2003, pp. 657-673.
- [3] C. Zhu, Y.-K. Wang, M.M. Hannuksela, and H. Li, "Error Resilient Video Coding Using Redundant Pictures," presented on *International Conference on Image Processing (ICIP)*, Atlanta, USA, Oct. 2006.
- [4] Y. Zhang, W. Gao, H. Sun, Q. Huang, and Y. Lu, "Error resilient video coding in H.264 encoder with potential distortion tracking," *International Conference on Image Processing (ICIP)*, vol. 1, pp.163-166, Oct. 2004.
- [5] Y. Wang, S. Wenger, J. Wen, and A.K. Katsaggelos, "Error resilient video coding techniques," *IEEE Signal Processing Magazine*, vol. 17, issue 4, pp. 61 - 82, Jul. 2000.
- [6] T. Wiegand and B. Girod, "Lagrangian multiplier selection in hybrid video coder control," in Proc. *International Conference on Image Processing (ICIP)*, Thessaloniki, Greece, October 2001.
- [7] Y. Guo, H.Q. Li, and Y.K. Wang, SVC/AVC loss simulator, *Joint Video Team*, Doc. JVT-Q069, Oct. 2005.