

H.26L의 블록화 및 링 현상을 동시에 제거하기 위한 저 계산량의 Loop 필터 방식

이민구, 차형태, 한현수, 홍민철
송실대학교 정보통신전자공학부

A Very Low Complexity Loop Filter to Simultaneously Reduce Blocking and Ringing Artifacts of H.26L Video Coder

Min Gu Lee, Hyung Tae Cha, Hern-Soo Hahn, and Min-Cheol Hong
School of Electronic Engineering, Soongsil University, Korea
mhong@e.ssu.ac.kr

요약

본 논문에서는 H.26L 동영상 압축 표준화 방식에서 블록화 및 링 현상을 동시에 제거하는 계산량이 감소된 일차원 loop filter를 제안한다. 새로운 일차원 정규화 완화 함수가 정의되고 두 개의 인접 방향에서 완화의 정도를 조절하는 정규화 매개변수는 부호화와 복호화 부에서 이용 가능한 코드화 블록 패턴과 양자화 스텝 크기로 정의된다. 그러므로, 정규화 매개변수를 정의하고 압축된 동영상으로부터 복원된 영상을 얻기 위한 기타 정보는 필요하지 않다. 실험결과로부터 제안된 알고리즘의 성능은 확인할 수 있었다.

1. 서론

ITU-T는 고화질의 동영상을 고압축율에서 얻기 위해 H.26L 동영상 압축 표준화를 개발하고 있다[1]. 기존의 8x8 블록 단위의 DCT 절차를 이용한 압축된 영상의 블록화 및 링 현상의 감소는 여러 논문을 통해 발표되어 왔다[2,3,4,5]. 하지만, H.26L은 4x4 블록 단위의 코딩과 공간 예측 방식을 사용하기에 코드화 계수들의 분포는 기존의 8x8 블록 단위 기반의 MC/DCT 방식들과 다르다. 비록 H.26L의 코딩 효율이 기존의 표준화된 압축 방식들에 비해 우수한 성능을 보인다. 하지만 더 적은 블록 크기로 인해 시각적으로 불편한 현상이 특별히 낮은 비트율에서 심각하게 나타난다. 그러므로, 만족스러운 화질의 압축된 동영상을 얻기 위해 효율적인 필터 기법이 요구된다[6].

TML5 (Test Model Long Term 5)에서 부호화와 복호화 부의 주요 함수들의 계산량을 보면 부호화 부의 움직임추정과 복호화 부의 loop filter가 주요 계산량인 것을 알 수 있다. 본 논문에서는 loop filter의 계산량을

줄이는데 초점을 둔다.

여러 동영상 압축 표준화 방식과 같이, 양자화는 TML5에서 원영상의 정보를 줄이는 주요 단계로 사용되었다. 그러나 비트율이 낮아짐에 따라, 현재 표준화 뿐만 아니라 존재하는 기법들의 대부분은 양자화 과정으로 인한 정보 손실로 복원영상에서 시각적으로 상당히 불편한 현상이 발생한다[1]. 양자화가 스텝 크기가 클수록 압축율은 높아지지만 정보 손실이 더 커지기 때문에 압축영상에 노이즈가 첨가되게 된다.

일반적으로 hybrid 동영상 압축 방식에서 발생하는 노이즈는 블록화 및 링 현상으로 분류할 수 있다. 블록화 현상은 영상을 블록 단위로 나누고 각 블록을 독립적으로 처리하여 저주파 성분의 정보 손실로 인한 현상이다. 반면에 링 현상은 고주파 성분의 정보 손실로 인해 비교적 단조로운 배경에 움직임이 있는 물체의 에지 부근을 따라 물결과 형태로 나타난다.

본 논문에서는 두 개의 인접 화소간에 완화를 부여하는 새로운 일차원 완화 함수를 해결책으로 정의한다. 제안된 함수는 지역적 데이터 신뢰도와 지역적 완화 제약을 결합한 두 개의 하위 함수들로 구성된다. 데이터에 지역적 신뢰도와 완화도 간의 균형을 조절하는 정규화 매개변수는 coded block pattern과 quantization step size와 같은 이용 가능한 부가 정보로 정의된다.

본 논문은 다음에 같이 구성된다. 2장에서는 새로운 일차원의 정규화 완화 함수가 나오고 그것의 특성들과 필터링 조건이 설명된다. 3장에서는 정규화 매개변수의 선택이 전형적인 정규화 이론에 따라 제안된다. 4장에서는 계산량의 부담을 줄이기 위한 수행 방식에 대해 설명된다. 마지막으로 실험 결과와 결론이 5, 6장에 이어진다.

2. 제안된 알고리즘

대부분의 동영상 압축 표준화에 따르면, 현재 H.26L 동영상 표준화 규격(TML)은 변환과 양자화 메커니즘을 사용한다. 양자화 과정으로 인한 정보 손실은 변환 영역에서 오기 때문에, 공간 영역에서 양자화 노이즈 형태 또는 분포를 추정이나 예측하는 것은 매우 복잡하다. 이를 간략히 하기 위해 양자화 노이즈를 i.i.d(identically independent distributed)라 가정하면, 화소 단위의 훼손된 모델은 다음과 같다.

$$g(i, j) = f(i, j) + n(i, j), \quad (1)$$

위 식에서 g , f , 및 n 은 압축영상, 원영상, 그리고 양자화 노이즈를 의미한다. 또한 (i, j) 는 이차원 영상을 표현하는 수직과 수평 방향 축을 의미한다.

새로운 loop filter 방식을 설명하기 이전에, coded block pattern이 본 논문에서 블록 strength를 정의하는데 어떻게 사용되는지 알아본다. 동영상 압축 표준화에서 계층적인 배열은 압축 효율을 최대화하는데 사용되어져 왔다. 그리고 각 레벨에서는 코딩 선택이 정의되고 부가 정보로 전송되어진다. 부가 정보 중의 하나인 coded block pattern은 intra-coded 블록, inter-coded 블록, not-coded 블록과 같은 각 블록의 코딩 형태를 포함한다. 이러한 정보는 코딩 형태마다 양자화 노이즈가 다른 분포를 가지고 화소간에 신호의 불균일의 정도가 다르기 때문에 완화도를 조절하는데 매우 유용할 수 있다.

시각적으로 불편한 현상은 블록이 intra-coded일 때, 인접 블록 간에 큰 움직임 벡터 차이를 갖을 때, 블록이 영이 아닌 변환계수를 가질 때 발견된다. 또한 블록 경계선에 화소의 신호의 불균일은 안쪽 화소의 신호의 불균일보다 더 높다. 이런 선행 정보는 완화의 정도가 블록 코딩 형태와 화소 위치에 따라 어떻게 선택되는지 미리 알 수 있다. 선행 정보를 기본으로 블록은 (1)intra-coded 블록(class1), (2)인접 블록간에 큰 움직임 벡터를 갖는 블록(class2), (3)영이 아닌 변환계수를 갖는 블록(class3)으로 분류된다. 각 블록에 블록 strength를 배정하는 것은 다음과 같다.

$$\text{Strength} = \begin{cases} 2 & \text{for class 1} \\ 1 & \text{for class 2} \\ 1 & \text{for class 3} \\ 0 & \text{otherwise} \end{cases}$$

위의 블록 strength, 화소 위치, 양자화 스텝 크기는 정규화 매개변수를 정의하는데 사용한다. 정규화 매개변수의 선택은 다음 장에서 설명되어진다.

H.26L에서는 코딩 메커니즘을 기본으로 하는 독립된 블록으로 압축영상의 불균일은 원영상의 불균일보다 더 크다. 그러므로 완화 제약은 시각적으로 불편한 현상을

줄이는데 필요로 한다. 정규화 이론은 영상복원 문제에 사용되어왔다. 전형적인 정규화 복원은 상당한 계산량을 필요로 하는 벡터-행렬 형태 및 반복 복원 기법을 사용한다. 매우 간단하면서 계산량 절감을 위한 방식을 얻기 위해서, 정규화 완화 함수를 기본으로 하는 일차원 화소는 다음과 같이 정의된다.(예를 들면, 수평방향)

$$M(f(i, j)) = M_p(f(i, j)) + M_n(f(i, j)) \quad (2)$$

여기서 각 정규화 완화 함수는 다음과 같다.

$$\begin{aligned} M_p(f(i, j)) &= (1 - \alpha_p(i, j))[f(i, j) - g(i, j)]^2 \\ &\quad + \alpha_p(i, j)[f(i, j) - f(i, j - 1)]^2, \\ M_n(f(i, j)) &= (1 - \alpha_n(i, j))[f(i, j) - g(i, j)]^2 \\ &\quad + \alpha_n(i, j)[f(i, j) - f(i, j + 1)]^2, \quad (3) \end{aligned}$$

위의 $\alpha_p(i, j)$ 와 $\alpha_n(i, j)$ 는 수평적으로 (i, j) 번째 화소의

이전과 다음 화소에서 완화의 정도를 조절하는 정규화 매개변수이고 0과 1 사이의 값을 갖는다. 각 정규화 매개변수는 상반되는 특성을 동시에 제한한다. 식 (2)의 해는 $f(i, j)$ 에 대해 미분하여 쉽게 얻을 수 있다. 또한 원영상의 두 인접하는 화소는 이용할 수 없기에 공간 영역에서 양자화 노이즈의 지역적 평균은 0과 같다고 가정한다.

$$\alpha_p(i, j)n(i, j - 1) + \alpha_n(i, j)n(i, j + 1) = 0. \quad (4)$$

위의 가정은 양자화 노이즈가 매우 작기에 높은 비트율에서 매우 합당하다. 그러나 비트율이 낮아짐에 따라 가정으로부터 편차는 더 높아지게 된다. 그럼에도 불구하고 방식을 쉽게 수행하게 하기에 가정을 식 (3)에 적용한다. 그러면 식 (3)의 해는 다음과 같다.

$$\begin{aligned} \hat{f}(i, j) &= \frac{(2 - \alpha_p(i, j) - \alpha_n(i, j))g(i, j) + A}{2}, \\ A &= \alpha_p(i, j)g(i, j - 1) + \alpha_n(i, j)g(i, j + 1). \quad (5) \end{aligned}$$

위의 정규화 완화 함수와 필터 과정은 수평방향으로 수행된다. 수직방향에 대한 필터 과정은 수평적으로 필터를 거친 영상과 같은 방법으로 대신한다.

3. 정규화 매개변수

정규화 매개변수의 선택은 완화의 정도를 조절하기에 정규화 복원 문제에 매우 중요한 것이다[3,4,5]. 앞장에서 기술된 바와 같이 전형적인 정규화 복원[4,5]에 따르면 정규화 매개변수는 원영상과 노이즈에 대한 선행 정보로 정의될 수 있다. 이러한 정보는 좋은 해결책을 달성하기 위해 이용 가능한 데이터와 부가 정보로부터 끌어낼 수 있다.

QP를 (i, j) 번째 화소를 포함한 매크로 블록의 양자화

스텝 크기라 하면 전형적인 정규화 이론[4,5]에 따라 정규화 매개변수는 다음과 같다.

$$\frac{1-\alpha_p(i,j)}{\alpha_p(i,j)} = \frac{[f(i,j)-f(i,j-1)]^2}{[f(i,j)-g(i,j)]^2} = \frac{[f(i,j)-f(i,j-1)]^2}{\phi(QP)},$$

$$\frac{1-\alpha_n(i,j)}{\alpha_n(i,j)} = \frac{[f(i,j)-f(i,j+1)]^2}{[f(i,j)-g(i,j)]^2} = \frac{[f(i,j)-f(i,j+1)]^2}{\phi(QP)}, \quad (6)$$

여기에서 $\phi(\cdot)$ 는 양자화 노이즈의 세기를 나타내기 위해 QP^2 에 비례해야 한다. 이전 장에서 언급되었던 것에 따라 선행 정보는 다른 완화의 정도는 다른 블록 분류와 화소의 위치를 가진다는 해결책으로 구체화할 수 있다. 완화도를 적응적으로 다루기 위해 tuning 매개변수를 정의하고 정규화 매개변수는 $n(i,j)=n(i,j-1)=n(i,j+1)$ 라는 가정 하에 다음과 같이 근사화 된다.

$$\alpha_p(i,j) = \frac{K_p QP^2}{[g(i,j)-g(i,j-1)]^2 + K_p QP^2},$$

$$\alpha_n(i,j) = \frac{K_n QP^2}{[g(i,j)-g(i,j+1)]^2 + K_n QP^2}, \quad (7)$$

tuning 매개변수는 블록 strength와 화소 위치에 의해 다음과 같이 정의된다.

$$K_p = \begin{cases} (2 * (m + 1)) / 16 & \text{for class 3} \\ (4 * (m + 1)) / 16 & \text{for class 2,} \\ (8 * (m + 1)) / 16 & \text{for class 1} \end{cases}$$

$$K_n = \begin{cases} (2 * (n + 1)) / 16 & \text{for class 3} \\ (4 * (n + 1)) / 16 & \text{for class 2,} \\ (8 * (n + 1)) / 16 & \text{for class 1} \end{cases} \quad (8)$$

여기서 m 과 n 은 화소 위치에 의해 다음과 같다.

$$m = \begin{cases} 1 & \text{if } g(i,j) \text{ and } g(i,j-1) \text{ are block boundary pixels} \\ 0 & \text{otherwise} \end{cases},$$

$$n = \begin{cases} 1 & \text{if } g(i,j) \text{ and } g(i,j+1) \text{ are block boundary pixels} \\ 0 & \text{otherwise} \end{cases}. \quad (9)$$

class 4 블록에 속하는 화소들에 대해 필터 과정은 실행되지 않는다.

4. 수행 방식

제안된 방식은 헤를 얻기 위해 floating-point 연산을 한다. floating-point 연산으로부터 오는 계산적 부담을 줄이기 위해 식 (5)와 integer 연산이 가능한 방식과 같이 상관하는 변형된 연산을 필요로 한다. 변형된 필터 과정은 다음과 같다.

$$\tilde{f}(i,j) = \frac{(2^9 - \beta_p(i,j) - \beta_n(i,j))g(i,j) + \beta_p(i,j)g(i,j-1) + \beta_n(i,j)g(i,j+1)}{2^9},$$

$$\beta_p(i,j) = 2^8 \alpha_p(i,j) = 2^8 \frac{K_p QP^2}{[g(i,j)-g(i,j-1)]^2 + K_p QP^2},$$

$$\beta_n(i,j) = 2^8 \alpha_n(i,j) = 2^8 \frac{K_n QP^2}{[g(i,j)-g(i,j+1)]^2 + K_n QP^2}. \quad (10)$$

제안된 방식에서 변형된 정규화 매개변수는 loop-up table과 같은 메모리에 저장하여 정규화 매개변수의

계산량을 줄인다. intensity 차에 대한 수, 양자화 스텝 크기, 화소의 위치는 한정되어 있기에 요구되는 메모리 공간은 거대하지 않다. 블록 형태 중 하나인 not-coded 블록은 현재 프레임의 코드화 블록과 관련 프레임의 코드화 블록 간에 높은 상관관계를 나타낸다. 필터 과정이 전형적인 블록화 현상을 줄인다 할지라도 그것은 계산량을 필요로 한다. 그러나 블록 pattern이 not-coded 블록으로 정의되었을 때, 계산량은 이전의 필터를 거친 프레임을 사용함으로써 계산량은 필요하지 않다. 그러므로 위의 정보를 이용한 제안된 필터 과정은 다음과 같다.

$$\tilde{f}(i,j) = \begin{cases} \tilde{f}_p(i,j) & \text{if } g(i,j) \in \text{not-coded block,} \\ \text{Eq.(10)} & \text{otherwise} \end{cases}, \quad (11)$$

여기서 $\tilde{f}_p(i,j)$ 은 복원된 이전 프레임의 (i,j) 번째 화소 강도를 나타낸다.

5. 실험 결과

제안된 loop filter는 H.26L TML5를 가지고 실험하였고 TML5 loop filter와 비교하였다. 제안된 방식의 성능을 측정하기 위해 PSNR(Peak to Signal to Noise Ratio)가 사용되었다. QCIF Foreman 영상은 10 frames/sec와 다양한 비트 율에서 실험하였다. de-blocking 필터를 사용하지 않은 97번째 압축된 프레임은 그림 1에서 보여진다. 또한 TML5 loop filter 및 제안된 loop filter를 사용한 복원한 영상은 그림 2,3에서 보여진다. 필터를 거친 결과는 블록화 현상이 잘 제거되어 있음을 알 수 있다. 그러나 TML5 loop filter를 사용해 복원한 영상은 특별히 non-stationary 움직임 물체 주위에 여전히 링 현상을 가지고 있다. 반면에 제안된 loop filter은 링 현상이 잘 제거되었다. 그림 4는 비트 율의 함수로써 TML5 base, TML5 loop filter, 제안된 loop filter의 100 frames의 평균 PSNR을 비교를 나타낸다. 제안된 loop filter 방식의 성능은 저 비트 율에서 TML5 loop filter보다 다소 향상되었다. 반면에 고 비트 율에서는 TML5 loop filter가 제안된 방식을 다소 능가한다.

제안된 방식의 이점은 계산량의 감소이다. 이는 그림 10에서와 같이 TML5 loop filter 계산량의 약 30~35%이다. 다양한 비트 율에서 제안된 방식의 계산량은 일관되게 감소한다.

6. 결론

H.26L 동영상 압축 표준화에서 블록화 및 링 현상을 동시에 제거하는 계산량이 감소된 loop filter를

제안한다. 새로운 완화 함수가 제안되었고 정규화 매개변수가 양자화 스텝 크기와 코드화 블록 패턴으로부터 효율적으로 정의되었다. 제안된 방식의 이점은 TML5 loop filter와 유사한 성능을 유지하면서 시각적으로 불편한 현상과 계산량을 효율적으로 줄였다는 것이다. 윤곽선을 보존하는 복원 영상을 얻기 위해 robust 함수에 대한 연구 중에 있다. 이 과정을 통해 더 정교하고 나은 성능을 발생하는 공식이 유도될 수 있을 것으로 사료된다.

참 고 문 헌

- [1] ITU-T SG16/Q15, ``H.26L Test Model Long Term Number 5 (TML-5),'' Aug. 2000.
- [2] R. Rosenholtz and A. Zakhor, ``Iterative Procedure for Reducing of Blocking Effects in Transform Image Coding,'’ *IEEE Trans. On Circuits and Systems for Video Technology*, vol. 2, pp. 91-94, Mar. 1992.
- [3] Y. Yang, N. P. Galatsanos, and A. K. Katsaggelos, ``Regularized Image Reconstruction from Incomplete Block Discrete Cosine Transform Compressed Images,'’ *IEEE Trans. On Circuits and Systems for Video Technology*, vol. 3, pp. 421-432, Dec. 1993.
- [4] M. R. Banham and A. K. Katsaggelos, ``Digital Image Restoration Algorithms,'’ *IEEE Signal Processing Magazine*, vol. 14, pp. 24-41, Mar. 1997.
- [5] M.-C. Hong, ``Adaptive Regularized Image and Video Restoration,'’ Ph.D Thesis, Northwestern University, Dept. of ECE, Dec. 1997.
- [6] M.-C. Hong, ``An Efficient Loop/Post Filter to Reduce Annoying Artifacts of H.26L Video Codec,'’ *IEEE Proceeding of International Conference on Consumer Electronics*, pp. 240-241, June 2000.



그림 1. Foreman 동영상의 de-blocking filter 없이 압축된 97th frame



그림 2. Foreman 동영상의 TML-5 loop filter로 복원된 97th frame



그림 3. Foreman 동영상의 제안된 loop filter로 복원된 97th frame

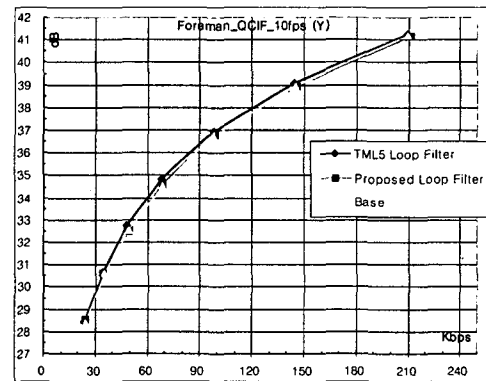


그림 4. QCIF Foreman 동영상의 비트 율에 따라 PSNR 비교

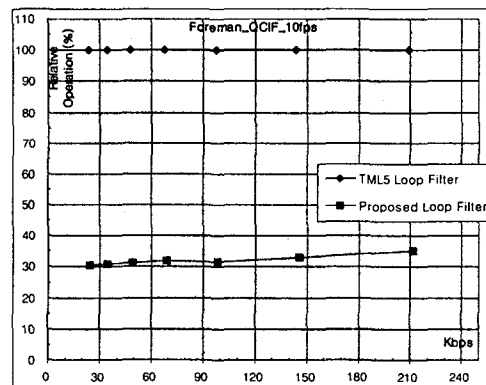


그림 5. QCIF Foreman 동영상의 비트 율에 따라 계산량 비교