

# 다중 레퍼런스 프레임 움직임 보상 기반의 효율적인 스케일러블 동영상 부호화 알고리즘

김승환\*, 김용관\*\*, 이상욱\*

\* 서울대학교 공과대학 전기공학부 뉴미디어통신공동연구소 영상처리연구실

\*\* 호서대학교 전기정보통신공학부

## Efficient Scalable Video Coding Based on Multiple Reference Frame Motion Compensation

Seung Hwan Kim\*, Yong Kwan Kim\*\*, and Sang-Uk Lee\*

\* Image Processing Lab., Inst. of New Media And Comm., SoEE, Seoul Nat'l Univ.

\*\* Information and Telecommunication Engineering, Hoseo Univ.

fax: +82-2-880-8220, email: cruise@ieee.org

### 요 약

기존 스케일러빌리티(scalability) 동영상 부호화 알고리즘의 문제점을 개선하기 위하여 본 논문에서는 다중 레퍼런스 프레임 방법을 기저 계층(base layer)과 확장 계층(enhancement layer)의 움직임 보상시 사용하여, 부호화 효율 및 드리프트 현상을 현저히 감소시키는 효과를 가져온다. 전산 모의 실험을 통하여 제안 알고리즘은 기존의 H.263+ 알고리즘에 비해 스케일러빌리티 모드를 작동하였을 때와 작동하지 않았을 경우 모두에 대해 더 나은 부호화 효율을 보여 주고, 이전 레퍼런스 확장 계층이 네트워크 사정으로 전송되지 않았을 경우에도 제안하는 알고리즘은 거의 화질의 손실이 없이 복원할 수 있음을 보여준다.

### I. 서 론

일반적인 인터넷 동영상 스트리밍(Internet video streaming) 시스템은 제한된 대역폭(bandwidth), 패킷 손실 비율(packet loss rate)등의 네트워크 성능 파라미터를 향상 고려하여 설계하여야한다. 이러한 인터넷 비디오 전송시 송신자와 수신자 간의 다양한 대역폭을 보상하기 위하여 그 해법으로 스케일러빌리티 비디오 부호화 알고리즘이 널리 사

용되고 있다 [2]. 더욱이 시간(temporal), 공간(spatial), 해상도(SNR) 스케일러빌리티 알고리즘은 MPEG-2, MPEG-4, H.263+[1]등과 같은 국제 비디오 부호화 표준안에 채택되어 사용되고 있다. 이러한 스케일러빌리티 알고리즘은 기저 계층(base layer)과 하나 이상의 확장 계층(enhancement layer) 구성된다. 기저 계층에는 복호기에서 스트림을 복원하는데 필요한 최소한의 정보가 저장되어 있고, 확장 계층은 복호기에서 영상의 화질을 개선하는 정보가 저장되어 있다.

그러나, 이러한 스케일러빌리티 비디오 알고리즘은 크게 두가지의 제약점이 있다. 첫째, 부호화 영상의 전송시 정보를 나누는 과정에서 총 데이터의 증가가 발생한다. 즉, 스케일러빌리티 비디오 부호화 알고리즘은 그것을 채택하지 않은 부호화 알고리즘에 비해 부호화 효율이 떨어진다. 둘째, 확장 계층의 정보의 손실은 다음 프레임의 부호화시 치명적인 손실을 가져온다. 이를 드리프트 효과(drift effect)라고 한다. 즉, 움직임 보상-이산 코사인 변환(MC-DCT) 알고리즘에 기반한 동영상 부호화 알고리즘은 움직임 보상 과정에서 참조 영상의 손실에 따른 치명적인 화질 저하를 가져온다.

이러한 문제점들을 개선하기 위하여, 여기서는 다중 레퍼런스 프레임 방법에 기반한 새로운 스케일러빌리티 동영상 부호화 알고리즘을 제안한다[3,4]. 움직임 보상의 효율을 향

상시키기 위하여 기저 계층(base layer)과 확장 계층(enhancement layer)의 움직임 보상을 할 때 다중 레퍼런스 프레임 방법을 제안한다. 또한, 이 다중 레퍼런스 프레임 방법을 이용하여 드리프트 효과를 현저히 감소시킬 수 있다.

본 논문은 다음과 같이 구성된다. II장에서는 제안하는 다중 레퍼런스 프레임 기반의 동영상 부호화 알고리즘에 대해 기술한다. III장에서는 제안하는 알고리즘의 전산 모의 실험 결과를 제시하고, IV장에서 본 논문의 결론을 짓는다.

## II. 제안 알고리즘

### 2.1 제안 움직임 보상 알고리즘

움직임 보상의 효율을 향상시키기 위하여, 기저 계층과 확장 계층 모두의 경우에 대해 다중 레퍼런스 프레임 기반의 움직임 보상 알고리즘을 제안한다. [4]에서 언급한 바와 같이 이중 레퍼런스 프레임(double reference picture) 방법과 이중 움직임 벡터(double motion vector) 방법은 움직임 보상의 효과와 복잡성간의 적절한 타협점이 된다. 그래서, 여기서는 기저 계층과 확장 계층 모두의 경우에 대해 이 두 가지, 이중 레퍼런스 프레임 방법과 이중 움직임 벡터 방법을 이용한다. 하지만, 기저 계층과 확장 계층은 그 신택스 구조(syntax structure)의 차이가 크므로, 똑같이 적용할 수는 없고 다른 알고리즘을 이용한다. 아래에서 R-D 최적화와 결합한 움직임 보상 방법에 대해 설명하겠다.

#### 2.1.1 기저 계층의 움직임 보상

기저 계층의 경우,  $k$  번째 프레임  $F_k$ 의  $i$  번째 블록  $B_i$ 의 움직임 보상을 할 경우, 두 개의 레퍼런스 프레임  $F_{k-1}^B$ 과  $F_{k-2}^B$ 로부터 두 개의 움직임 벡터  $v_{i,1}$ 과  $v_{i,2}$ 를 얻을 수 있다. 여기서 움직임 벡터  $v_{i,1}$ 에 의해 보상되는 블록을  $\hat{B}_{i,1}$ 라 하고, 움직임 벡터  $v_{i,2}$ 에 의해 보상되는 블록을  $\hat{B}_{i,2}$ 라 하면 블록  $B_i$ 는 다음과 같이 움직임 보상이 된다.

$$\hat{B}_i = h_1 \hat{B}_{i,1} + h_2 \hat{B}_{i,2}, \quad (2.1)$$

여기에서  $h_1$ 과  $h_2$ 는 고정된 계수 값이고,  $h_1 + h_2$ 는 정

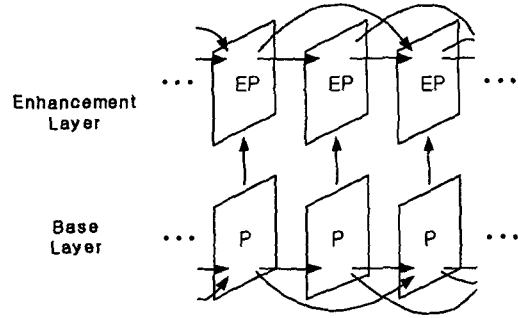


그림 1: 제안 스케일러빌리티 동영상 부호화 알고리즘

규화를 위하여 1로 둔다.

여기에서 기저 계층의 움직임 예측을 할 때, 효율을 증가시키기 위하여 R-D 최적화를 수행한다. 즉, 움직임 벡터  $v_{i,1}$ 과  $v_{i,2}$ 을 예측할 때, Lagrangian 비용 함수  $J = D + \lambda R$ 을 최소화 하는 값을 예측한다. 여기에서는 움직임 보상 블록  $\hat{B}_i$ 은 아래의 비용함수를 최소화하는 움직임 벡터를 결정하여 사용된다.

$$\arg \min_v [ \|B_k - (h_1 \hat{B}_{k,1} + h_2 \hat{B}_{k,2})\| + \lambda (R_{motion} + R_{DCT}) ], \quad (2.2)$$

여기에서, 움직임 정보의 양(rate)을 의미하는  $R_{motion}$ 은 단순히 움직임 벡터 표로부터 계산할 수 있지만, 이산 코사인 변환의 양을 의미하는  $R_{DCT}$ 는 동작 R-D 곡선을 예측하는데 발생하는 큰 계산량 때문에 얻기가 쉽지 않다. 그래서, 우리는 [5]에서 제안된 양의 모델을 이용하고 이는 아래와 같다.

$$R_{DCT} = \frac{1}{16^2} \left( K - \frac{\sigma_i^2}{Q_i^2} \right), \quad (2.3)$$

여기에서,  $\sigma_i$ 는 블록의 실험적인 표준 편차(standard deviation)의 값이고,  $Q_i$ 는 그 블록의 양자화(quantization) 간격이다.  $K$ 는 이산 코사인 변환의 계수가 Laplacian 분포를 가지며 독립이라면,  $e/\ln 2$ 이다.

#### 2.1.2 확장 계층의 움직임 보상

확장 계층의 경우,  $i$  번째 블록  $E_i$ 의 움직임 보상을 할 경우, 두 개의 레퍼런스 프레임  $F_{k-1}^E$ ,  $F_{k-2}^E$ 과 기저 계층  $\hat{B}_i$ 로부터 세 개의 움직임 벡터  $v_{i,1}$ ,  $v_{i,2}$ 과  $v_{i,b}$

를 얻을 수 있다. 그림 1처럼  $v_{i,b}$ 는 시간상으로 동일한 기저 계층으로부터 움직임 보상을 했을 경우의 움직임 벡터이며 (0,0)의 값을 가진다. 또한,  $v_{i,1}$ 은  $\tilde{F}_{k-1}^E$  레퍼런스 프레임으로부터 움직임 보상을 하고,  $v_{i,2}$ 는  $\tilde{F}_{k-2}^E$ 로부터 움직임 보상을 한다. 기저 계층과는 달리 확장 계층에는 각각의 움직임 벡터의 제한을 두어 서로 다른 레퍼런스 프레임에서만 움직임 보상을 한다. 확장 계층의 블록  $E_i$ 는 다음과 같이 움직임 보상이 된다.

$$\hat{E}_i = h_1 \hat{E}_{i,b} + h_2 \hat{E}_{i,1} + h_3 \hat{E}_{i,2}, \quad (2.4)$$

여기에서  $h_1$ ,  $h_2$ ,  $h_3$ 는 고정된 계수 값이고,  $h_1 + h_2 + h_3$ 의 합은 정규화를 위하여 1로 한다.

여기에서 확장 계층의 움직임 예측을 할 때 또한, 효율을 증가시키고, 실시간 전송을 위한 버퍼(buffer)에 저장된 데이터의 양을 조절하기 위하여 R-D 최적화를 수행한다. 하지만, 기저 계층에서는 매크로 블록 별로 양자화 간격을 조금 조절할 수 있으나, 확장 계층의 R-D 최적화 과정에서는 양자화 간격이 프레임별로 고정된다. 움직임 벡터  $v_{i,1}$ 과  $v_{i,2}$ 을 예측할 때, Lagrangian 비용 함수  $J = D + \lambda R$ 을 최소화 하는 값을 예측한다. 움직임 보상 블록  $\hat{E}_i$ 은 아래의 비용함수를 최소화하는 움직임 벡터를 결정하여 사용된다.

$$\arg \min_v [ \|E_i - (h_1 \hat{E}_{i,b} + h_2 \hat{E}_{i,1} + h_3 \hat{E}_{i,2})\| + \lambda (R_{motion} + R_{DCI}) ]. \quad (2.5)$$

## 2.2 드리프트 효과

이전 프레임의 확장 계층이 네트워크의 혼잡(network congestion) 때문에 전송되지 않았을 경우, 레퍼런스 프레임이 없으므로, 이후의 프레임은 정상적으로 움직임 보상을 할 수 없으므로, 에러 전이(error propagation)이 발생한다. 이러한 문제를 해결하기 위하여, 이전의 확장 계층으로부터가 아닌 이전의 기저 계층으로부터 움직임 보상을 하는 알고리즘들이 제안되었다 [5]. 이것은 비록 드리프트 현상이 없어지기는 하지만, 부호화 효율이 현저히 떨어지는 단점이 있다. 그러므로, 드리프트 현상과 부호화 효율을 동시에 해결하기 방법의 해가 다중 레퍼런스 프레임을 이용한 움직임 보상이다.

(2.4)에 언급한 것처럼 제안하는 알고리즘의  $i$ 번째 확장 계층의 블록  $E_i$ 의 움직임 보상은  $\hat{E}_{i,b}$ ,  $\hat{E}_{i,1}$ ,  $\hat{E}_{i,2}$ 로부터 중첩의 원리를 이용한다. 여기에서,  $\hat{E}_{i,b}$ ,  $\hat{E}_{i,1}$ ,  $\hat{E}_{i,2}$ 는 모두 시간적으로 같은 영상  $F_k$ 를 움직임 보상한 것이므로, 서로 간에 큰 상관도(correlation)를 가진다. 그러므로, 레퍼런스 프레임  $\tilde{F}_{k-1}^E$  또는  $\tilde{F}_{k-2}^E$ 이 전송되지 않았을 경우, 복호기는 에러 전이를 감소시키며 효율적인 복원을 할 수 있다. 즉, 만약  $\tilde{F}_{k-1}(u, v)$ 가 손실되었다면, 복원된 화소  $\hat{E}_i^d(u, v)$ 는 다른 두 블록을 이용하여 다음과 같이 복원한다.

$$\hat{E}_i^d(u, v) = \frac{h_1}{h_1 + h_3} \hat{E}_{i,b}(u, v) + \frac{h_3}{h_1 + h_3} \hat{E}_{i,2}(u, v), \quad (2.6)$$

그리고, 만약  $\tilde{F}_{k-2}(u, v)$ 가 손실되었다면, 복원된 화소

$$\hat{E}_i^d(u, v)$$

$$\hat{E}_i^d(u, v) = \frac{h_1}{h_1 + h_2} \hat{E}_{i,b}(u, v) + \frac{h_2}{h_1 + h_2} \hat{E}_{i,1}(u, v) \quad (2.7)$$

로 구한다. 마지막으로, 아무런 손실이 없다면

$$\hat{E}_i^d(u, v) = h_1 \hat{E}_{i,b}(u, v) + h_2 \hat{E}_{i,1}(u, v) + h_3 \hat{E}_{i,2}(u, v), \quad (2.8)$$

로 구한다.

이러한 방법을 통하여 복원된 영상은 드리프트 문제가 없이 완벽하게 복원은 되지 않지만, 에러를 현저히 감소시키며 좋은 화질로 복원을 할 수 있다. 이것은 [3]에 제시된 분석을 조금 변형하면,  $F_k$  프레임이 전송 중에 손실되었을 경우,  $(k+1)$ 번째 프레임의 에러의 한계점  $\|e_{k+1}\|$ 은

$$\|e_{k+1}\| \leq \max_m (\| \hat{E}_{m,1} - (\hat{E}_{m,2} + \hat{E}_{m,b}) \|) \frac{1 + h_2 + h_2^2}{2 - h_2} \left[ 1 - \frac{(h_2 - 1)^k}{1 + h_2 - h_2^2} \right], \\ \|e_{k+1}\| \leq \max_m (\| \hat{E}_{m,1} - (\hat{E}_{m,2} + \hat{E}_{m,b}) \|) \frac{1 + h_2 + h_2^2}{2 - h_2} \quad (2.9)$$

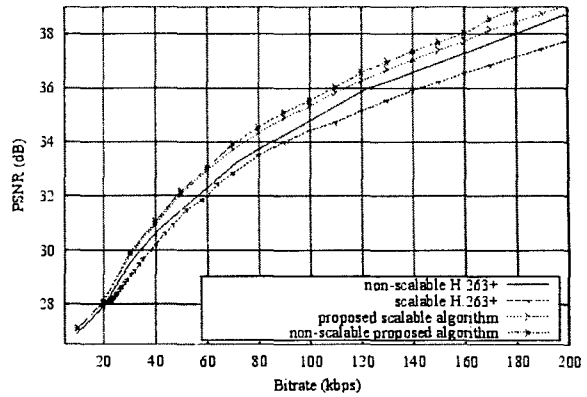
이다. (2.9)에서, 에러 한계점은  $h_2$ 에 관한 식으로 표현된다. 즉,  $h_2$ 가 증가할수록 에러의 한계점이 증가하고,  $h_2$ 가 감소할수록 에러의 한계점은 0에 근접한다. 따라서,  $h_2$ 의 값을 증가시킬수록 에러에 강인한 결과를 가져올 수 있다. 하지만,  $h_2$ 의 값을 감소시키면, 역으로 부호화 효율이 감소되는 결과를 가져오므로, 부호화 효율과 에러의 강인성을 동시에 고려하여 적절한  $h_2$ 의 선택이 필요하다.

### III. 전산 모의 실험

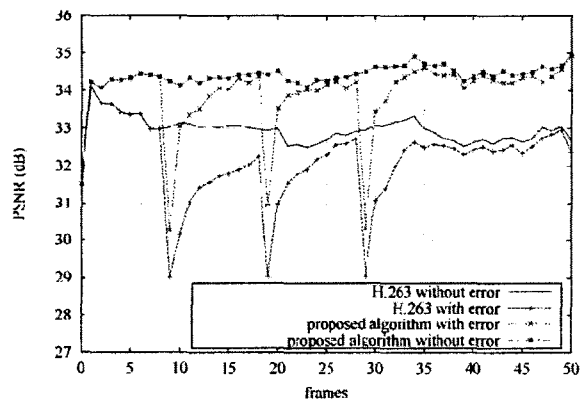
제안하는 알고리즘들은 H.263+ TMN 8.0 동영상 표준안의 신택스를 수정하여 구현하였다. 실험 영상은 표준 QCIF Foreman 300 프레임 시퀀스를 사용하였다. 그림 2(a)는 부호화 효율을 비교하기 위해, 스케일러빌리티 모드를 작동시킨 것과 작동시키지 않은 두가지 경우에 대해, 제안하는 알고리즘과 H.263+ 알고리즘을 비교하였다. H.263+ 알고리즘의 경우 각각의 비트양에 대해 곡선을 그리기 위하여 Annex T의 R-D 최적화 알고리즘을 동작시켰다 [1]. 제안 알고리즘과 H.263+의 스케일러빌리티 모드의 경우, 기저 계층의 비트양을 20 kbps로 고정 시키고, 확장 계층의 비트양은 QUANT를 조절하여 그래프를 생성하였다. 그림으로부터 제안 스케일러빌리티 알고리즘들은 60 kbps에서 H.263+ 알고리즘에 비해 1 dB정도 우수한 성능을 보여 줌을 알 수 있다. 30 kbps이하의 낮은 비트율에서도 제안 알고리즘들은 H.263+의 두 가지 모드에 대해 모두 우수한 성능을 보여줌을 알 수 있다. 그림 2(b)는 확장 계층의 10번째, 20번째, 30번째 프레임이 네트워크 사정으로 전송되지 않았을 경우의 PSNR 성능을 보여준다. 기본 H.263+ 부호기는 확장 계층이 전송되지 않았을 경우, 치명적인 화질의 손실이 있고, 그 에러가 계속해서 전이되어 그 이후의 프레임에도 영향을 줌을 볼 수 있다. 하지만, 제안 알고리즘의 경우, 에러를 효과적으로 국소화(localization)하여, 에러가 발생한 이후의 프레임에는 거의 영향을 받지 않고, 에러가 없는 경우와 거의 유사한 화질을 보여준다.

### IV. 결론

본 논문에서는 부호화 효율을 높이고 드리프트 현상을 감소시키기 위하여 다중 레퍼런스 프레임 방법을 이용하여 기저 계층과 확장 계층의 움직임 보상을 하는 알고리즘을 제안하였다. 실험을 통하여 제안 알고리즘들은 H.263+에 비해 스케일러빌리티 모드를 작동하였을때와 작동하지 않았을 경우 모두에 대해 더 나은 부호화 효율을 보여 주고, 이전 레퍼런스 확장 계층이 전송되지 않았을 경우에도 제안하는 알고리즘들은 거의 화질의 손실이 없음을 보여주었다.



(a) 부호화 효율



(b) 드리프트 효과

그림 2. PSNR 성능의 비교: QCIF Foreman 시퀀스

### 참고 문헌

- [1] *Video Coding for low bitrate communication*, Jan. 1998. ITU-T recommendation H.263; version 2.
- [2] H. Radha, Y. Chen, K. Parthasarathy, and R. Cohen, "Scalable Internet video using MPEG-4," *Signal Processing: Image communication*, vol. 15, pp. 95-126, 1999.
- [3] C.-S. Kim, R.-C. Kim, and S.-U. Lee, "Robust transmission of video sequence using double-vector motion compensation," accepted in *IEEE Trans. Circuits Syst. video Technol.* (in press)
- [4] S. H. Kim, C.-S. Kim, and Sang-Uk Lee, "Efficient motion compensation algorithm based on second-order prediction," in *Proc. IEEE Int. conf. Image Processing*, vol. 2, pp. 875-878, Sept. 2000.
- [5] W. Li, "Overview of fine granularity scalability in MPEG-4," in *IEEE Trans. Circuits Syst. video Technol.*, vol. 11, no. 3, pp. 301-317, March 2001.