

웨이블릿 코덱에서의 Inter-Intra 움직임 예측 기법¹⁾

*이주경[○] *김충길 **강정구 *정기동

*부산대학교 전자계산학과

** (주) 미디어트랜스

{jkleee[○], chgkim, jgkang, kdchung}@melon.cs.pusan.ac.kr

Inter-Intra Motion Estimation in Wavelet based Codec

*Joo-Kyong Lee[○] *Chung-gil Kim **Jeong-Gu Kang *Ki-Dong Chung

*Dept. of Computer Science, Pusan National University

** Media Trans Technology Inc.

요 약

웨이블릿 변환에 기반한 동영상 코덱에서의 움직임 예측 기법은 DCT 기반 코덱과 유사하게 이전 프레임과의 움직임 예측을 통하여 수행된다. 그러나, 현재 프레임이 이전 프레임에 참조하므로 네트워크상의 전송시 이전 프레임에 발생한 오류가 전달되는 오류 전파의 문제도 발생하게 된다. 본 논문에서는 웨이블릿 변환된 프레임의 특성을 이용하여 최상위 레벨의 LL 부대역만 이전 프레임과의 움직임 예측을 수행하고, 나머지 부대역에 대하여 프레임 내의 상위레벨의 부대역이 하위 부대역을 참조하여 예측 및 보상을 수행하여 오류전파의 가능성을 최소화하는 Inter-Intra ME 동영상 코덱을 제안한다. 제안된 움직임 예측을 사용하여 MAD(Mean-Absolute Differences)를 측정할 결과, 프레임간 변화가 심한 경우에는 제안된 기법과 이전 프레임의 부대역을 참조한 기법 사이의 압축율은 유사하게 나타났으며, 변화가 적은 경우에는 이전 프레임을 참조하는 것의 압축율이 높게 나타났다. 그러나, 네트워크 전송시 발생하는 오류전파에는 제안된 기법의 성능이 우수한 것으로 나타났다.

1. 서 론

웨이블릿 변환은 다양한 분야에서 적용되고 있으며 특히 이미지 처리 분야에서는 그 성능이 인정되어 JPEG-2000[1] 표준에 채택된 바 있다. 웨이블릿 변환이 이미지와 비디오 코딩에 인기있는 이유는 기본적으로 변화가 있는 비디오 데이터를 효율적으로 표현할 수 있으며, 이웃한 데이터 간의 유사성 제거와 전체 에너지의 효율적인 이동 및 압축이 가능하고, 이미지의 질을 떨어뜨리는 블로킹 효과가 없다는 점이다[2]. 웨이블릿 변환된 이미지는 서로 다른 주파수 영역의 다해상도 부대역으로 분할되며 각 부대역의 웨이블릿 계수의 특성을 이용한 압축 기법이 연구되어 왔다[3],[4]. 비디오 코딩에서는 프레임간 중복성을 줄이기 위한 다양한 종류의 기법[5]이 연구되어 왔으며 특히, 웨이블릿 도메인 상에서의 움직임 예측을 통한 보상기법이 많이 연구되었다. 이 기법들은 웨이블릿 도메인에서의 움직임 예측의 복잡도를 줄이거나[2],[6] Shift-invariant 특성 때문에 발생하는 성능저하를 줄이기 위한 방법[7],[8]을 제안했다. 이 기법들은 이전 프레임을 참조하여 움직임 예측을 수행한다는 측면에서는 동일하다.

본 논문에서는 웨이블릿 도메인에서 최상위 레벨의 LL 부대역을 제외한 나머지 부대역은 현재 프레임 내에서 움직임 예측 및 보상을 수행하고, 최상위 레벨의 LL 부

대역만 이전 프레임을 참조하여 움직임 예측을 수행하는 예측 기법을 제안한다. 이 기법은 이전 프레임에 대한 의존도를 최소화하여 오류전파의 가능성을 줄이고, 압축 효율은 기존의 움직임 예측 기법의 성능을 유지하도록 한다.

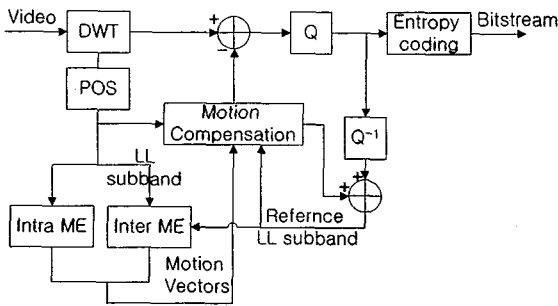
2. Inter-Intra 움직임 예측

2.1 부호기의 블록도

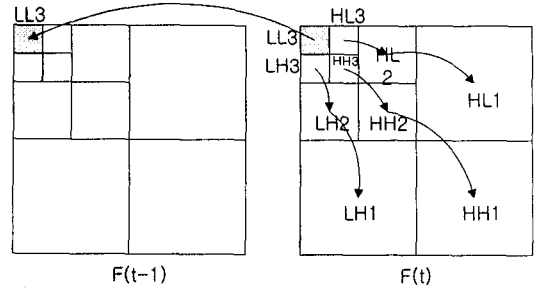
본 논문에서 제안하는 Inter-Intra 웨이블릿 부호기의 블록도는 (그림 1)과 같다. 제안하는 기법에서는 입력되는 비디오 프레임에 대하여 DWT(Discrete Wavelet Transform)를 수행하여 움직임 예측의 입력으로 사용한다. 움직임 예측을 수행하기 전, 웨이블릿 계수는 POS블록에서 양수로 변환되는 단계를 거친다. 일반적으로 웨이블릿 변환된 이미지의 계수값은 소스 이미지의 값과 변환에 사용된 필터의 종류에 따라 다르다. 그러나, 대부분 최상위 레벨의 LL부대역의 계수값은 에너지가 높은 양수의 값으로, 나머지 부대역의 계수는 양수 및 음수 계수가 섞인 형태로 나타난다. 이 같은 특징은 이전 프레임과의 움직임 예측을 수행하거나 프레임내의 다른 레벨의 부대역과의 움직임 예측을 수행할 때 효율을 떨어뜨리는 요인이 된다. 그러므로 본 논문에서는 웨이블릿 계수를 양수로 변환하는 단계를 거쳐 움직임 예측에 이용한다. 본 논문과 기존의 웨이블릿 코덱과의 가장 큰 차이점이 움직임 예측을 수행하는 블록이 Intra와 Inter 영역으로 나누어져 있는 점이다. Intra ME 블록에서는 최상위의 LL 부대역을 제외한 부대역에 대해 Intra ME를 수행하며, Inter ME 블록에서는 복원된 이전 프레임의

¹⁾ 본 연구는 (주) 미디어트랜스의 지원으로 수행되었음.

최상위 LL을 참조하여 움직임 예측을 수행한다. 움직임 보상을 수행하는 블록의 입력은 세가지가 된다. 즉, Inter ME와 Intra ME 블록에서 입력받는 모션벡터, 움직임 보상에 이용할 DWT 변환된 현재 프레임과 이전 프레임의 최상위 LL 부대역이다. 움직임 보상이 수행된 데이터는 양자화되고, 엔트로피 코딩을 거치게 된다. 양자화된 데이터는 다음 프레임에서 참조할 수 있도록 역양자화 단계 및 역움직임 보상 단계를 거쳐 메모리에 저장된다. 이때, 다음 프레임에서 참조하는 영역은 최상위 레벨의 LL 부대역이므로 이 영역만 역양자화, 역움직임보상 단계를 수행하여 저장함으로써 계산에 사용되는 연산수와 메모리를 절약할 수 있다.



(그림 1) Inter-Intra 움직임 예측을 수행하는 웨이블릿 부호기의 블록도



→ 참조 방향
 □ 최상위 레벨의 LL 부대역
 □ 움직임 예측에 이용되는 부대역

(그림 2) 3레벨 웨이블릿 변환 프레임의 움직임 예측 예

$$SAD(i, j) = \sum_{k=0}^{px-1} \sum_{l=0}^{py-1} \{ |S_M(x+k, y+l) - K \times S_{M-1}(2x+k+i, 2y+l+j)| \} \quad (1)$$

$$MV = \arg \min SAD(i, j), -p \leq i, j \leq p$$

M : 최상위 웨이블릿 변환 레벨

$S = \{LH, HL, HH\}$

K : 레벨 간 계수의 차이 보전을 위한 계수

px, py : 최상위 레벨 블록의 가로 및 세로 크기

p : search area

나머지 레벨에서의 움직임 예측에는 직상위 레벨의 모션벡터가 이용된다. 즉, 직상위 레벨 해당 블록의 모션벡터를 가로, 세로 2배씩 곱한 후, 참조영역의 위치로 이동한 후 움직임 예측을 수행한다. 이때, 움직임 예측 기법은 (식 1)과 동일하다. 다만, 움직임 예측 수행 범위는 $p \times 2^{M-m}$ 로 줄어든다. 이 기법은 MRME[6]과 동일하다.

$$MV_m = MV_{m+1} \cdot 2^{M-m-1} + \Delta(\delta x, \delta y) \quad (2)$$

m : 레벨, $1 < m < M$

움직임 보상 단계는 계산된 모션벡터를 이용하여 참조 프레임과 현재 프레임의 차이값을 저장하게 된다.

$$S_m = S_m - MV_m \cdot S_{m-1} \quad (3)$$

2.2.2 Inter ME · MC

최상위 레벨의 LL 부대역의 움직임 예측 및 보상은 기존의 코덱에서의 움직임 예측과 동일한 방법으로 수행한다. 다만, 움직임 예측 범위가 이전 프레임의 LL 부대역으로 한정된다는 점만 다르다.

2.2 움직임 예측 및 보상 처리과정

2.2.1 Intra ME · MC

(그림 2)는 3레벨 웨이블릿 변환된 프레임에 대해 Intra ME 및 Inter ME 블록에서 수행하는 움직임 예측의 예를 보여준다. $F(t)$ 는 현재 프레임을 의미하고, $F(t-1)$ 은 직전 프레임을 의미한다. 웨이블릿 변환된 이미지는 주파수 영역이 다른 다해상도의 부대역으로 분할되는 특징이 있다. 즉, 레벨에 따라 부대역의 크기가 다르며, 픽셀당 계수의 크기도 다르다. 또한 상위레벨과 하위레벨의 부대역간의 유사성도 존재한다. 예를 들어 LH3은 LH2의 1/4 크기이며 형태는 LH2를 축소한 형태로 나타나며 계수당 크기는 K배 높은 것으로 나타난다. K는 소스이미지, 필터에 따라 다르게 나타난다. 웨이블릿 변환 이미지의 이러한 특징은 Intra ME를 수행할 수 있는 조건을 만족시켜준다. (그림 2)에서 보듯이 Intra ME는 상위레벨의 부대역이 하위 부대역을 참조하여 움직임 예측을 수행한다. 따라서, 최하위 레벨의 부대역은 LH1, HL1, HH1은 움직임 보상이 수행되지 않는다. 이 부대역은 양자화와 엔트로피 코딩에만 이용된다. Intra ME를 위해 전체 부대역을 동일한 수의 블록으로 분할한다. 그러므로 하위레벨은 상위 레벨 블록의 4배의 크기가 된다. (식 1)은 최상위 레벨 M 의 (x, y) 번째 블록에 대한 모션벡터를 찾는 과정을 기술한 것이다. SAD는 현재 블록내의 각 픽셀과 참조 영역의 픽셀간 차이값의 절대치의 합을 구하는 함수이다. 계산된 함수 값 중에서 가장 값이 작은 함수의 위치 (i, j) 가 모션벡터가 된다.

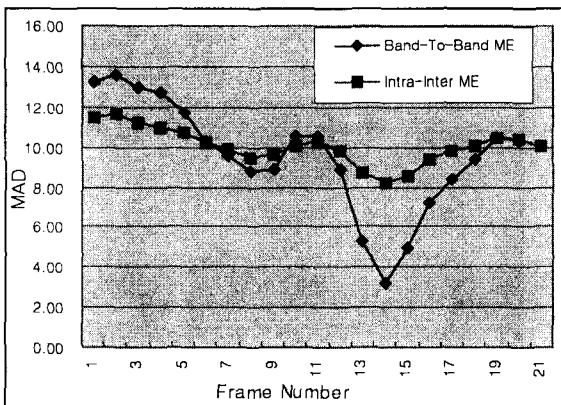
3. 실험결과

2.2.1에서 밝혔듯이 상위 부대역의 픽셀당 계수값은 하위 부대역의 계수값 보다 K배 크다. Intra ME를 효율적으로 수행하기 위해서는 이 K 상수를 적절하게 선택하는 것이 매우 중요하다. 본 연구에서는 화면의 변화가 적은 비디오와 변화가 심한 비디오를 다양하게 분석하여 움직임 예측에 사용할 상수 K를 선택하기로 한다. 실험에 사용된 이미지는 QCIF(176×144) 포맷으로 Akiyo, Foreman, Container, Carphone, Silent 등이며 웨이블릿 변환에 사용된 필터는 7/9 Antonini를 사용하였으며 웨이블릿 변환 레벨은 3레벨로 실험하였다. (표 1)은 실험에 사용된 동영상의 평균값을 나타낸 것이다. 비율이라는 것은 상위레벨의 픽셀당 계수값 / 하위레벨의 픽셀당 계수 값을 의미한다. 대체로 상위에서 하위로 내려갈수록 비율이 높아지며, 같은 레벨의 부대역간에도 비율이 다를 수 있다. 또한, 비디오의 화면 변화의 정도보다 이미지의 형태에 따라 값이 다를 수 있었다. 실험결과 LH 부대역의 K = 3이 적당하며, HL 영역의 경우 K = 2.5 ~ 3.5, HH의 영역의 경우 2.5 ~ 5로 주는 것이 적당함을 알 수 있었다.

(표 1)

부대역	비율(K)	부대역	비율(K)	부대역	비율
LH3	3.0	HL3	2.5	HH3	2.6
LH2	3.3	HL2	3.6	HH2	5.0
LH1		HL1		HH1	

(그림 3)은 웨이블릿 도메인상에서 이전 프레임의 부대역을 참조하는 band-to-band 움직임 예측과 제안된 기법의 성능을 평균 MAD(Mean Absolute Differences)로 나타낸 것이다. 사용된 비디오는 움직임의 변화가 심한 Foreman 데이터이다.



(그림 3) Foreman 동영상에 대한 평균 MAD값 비교

실험을 통해 제안된 기법의 MAD값은 변화가 적은 반면 이전 프레임과의 움직임 예측을 수행하는 경우에는 이전 프레임과 현재 프레임의 변화 정도에 따라 MAD값의 차이가 심하다는 것을 알 수 있다. 또한, Akiyo와 같

이 화면변화가 매우 적은 경우에는 이전 프레임과의 움직임 예측이 유리함을 알 수 있었다. 그러나, 제안된 기법은 이전 프레임에 대한 의존성이 낮아서 패킷 손실에 따른 오류전파의 가능성은 낮아서 수신측에서의 비디오 품질은 높은 편임을 알 수 있었다.

4. 결론 및 향후 연구과제

본 논문에서는 웨이블릿 도메인상에서 움직임 예측을 수행할 때, 이전 프레임에 대한 의존성을 줄이고, 현재 프레임 내에서 움직임 예측을 수행하여 네트워크 상으로 패킷 전송시 발생하는 오류전파를 최소화하는 Inter-Intra 움직임 예측 기법을 제안하였다. 제안된 기법은 프레임간의 움직임 변화가 심한 비디오에서는 기존의 Band-to-Band 기법과 유사한 압축 성능을 보였으며, 프레임간 변화가 적은 비디오에 대해서는 낮은 성능을 보였다. 그러나, 이전 프레임에 대한 참조율이 낮으므로 오류전파의 가능성을 줄어드는 장점이 있다. 향후 연구과제로는 인터넷 또는 모바일 상에서의 오류전파 실험을 통하여 제안된 기법의 성능을 측정하는 것이다.

5. 참고 문헌

- [1] <http://jpeg.org/CDs15444.html>, ISO/IEC 15444-1:2000
- [2] M.K. Mandal, E.Chan, X. Wang and S. Panchanathan, "Multiresolution Motion Estimation Techniques for Video Compression," Optical Engineering, Vol. 35, No. 1, pp.128-136, Jan. 1996.
- [3] J. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," IEEE Trans. Signal Processing, vol.41, pp.3445-3462, Dec. 1993.
- [4] A. Said and W. Pearlman, "A new, fast, and efficient imae codec based on set partitioning in hierarchical trees," IEEE Trans. Circuits Syst. Video Technol. vol. 6, pp.243-250, June 1996.
- [5] D. Taubman and A. Zakhor, "Multirate 3-D subband coding of video," IEEE Trans. Image Processing, vol. 3, pp.572-588, Sept. 1994.
- [6] Y. Zhang and S. Zafar, "Motion-compensated wavelet transform coding for color video compression," IEEE Trans. Circuits Syst. Video Technol., vol.2 pp.285-296, Sept. 1992
- [7] S. Kim, S. Ree, J. G. Jeon, and K. T. Park, "Interframe coding using two-stage variable block-size multiresolution motion estimation and wavelet decomposition," IEEE Trans. Circuits Syst. Video Technol. vol. 8, pp.399-410, Aug. 1998.
- [8] H. W. Park, H. S. Kim, "Motion Estimation Using Low-Band-Shift Method for Wavelet-Based Moving-Picture Coding," IEEE Trans. on Image Processing, Vol. 9(4), pp577-587, 2000