

3 차원 웨이블릿 기반의 스케일러블 비디오 부호화를 위한 효과적인 움직임 예측 및 보상 방법

김종호*, 이준재, 정제창
한양대학교 전자통신전파공학과

Efficient motion estimation and compensation methods for scalable video coding using 3D wavelet transform

Jongho Kim*, Joon-Jae Lee, Jechang Jeong
Dept. Electrical and Computer Engineering, Hanyang University
E-mail : {*angel, blue1336, jjeong}@ece.hanyang.ac.kr

Abstract

In this paper, the efficient motion estimation and compensation method for 3 dimensional wavelet transform is proposed. Recently, since the compression performance and scalable functionality are provided by wavelet transform, many researches have been carried out for applying to the video compression. For the temporal filtering, motion estimation and compensation techniques are used, but the unconnected pixels, which are produced by motion compensation result into the degradation of coding performance and quality of the picture. For the efficient motion compensated temporal filtering by reducing the number of these unconnected pixels, we propose the variable block size motion estimation and compensation method. Also we propose a method that determines the block size using rate-distortion optimization technique according to the local characteristics of the frame. The simulation results show the improved performances than the MPEG-4 scalable coding methods and the 3 dimensional wavelet coding methods using fixed block size motion estimation and compensation.

I. 서론

멀티미디어 기술의 발전과 MPEG을 비롯한 고효율의 국제 표준들의 완성과 더불어 휴대폰 및 PDA 등과 같이 다양한 단말을 이용한 컨텐츠의 소비요구가 급속도로 증가하고 있고 상대적으로 안정적인 유선망 뿐만 아니라 무선망을 이용한 서비스 요구도 급증하고 있다.

이러한 요구의 충족을 위해 전통적으로 트랜스코딩 및 스케일러블 방식의 부호화 방법이 고려되어 왔다.

트랜스코딩 방식은 이미 부호화되어 있는 비트스트림을 목표 시스템에 적합한 형태로 변환하는 방식이지만 이에 따르는 화질열화 및 지연시간의 문제가 발생한다. 반면 스케일러블 방식은 비트스트림을 계층(layer) 구조로 부호화하여 다양한 단말 및 네트워크 상황에 적절한 정도까지만 복호할 수 있도록 하는 방식이다. 하지만 기존의 방식들은 네트워크의 트래픽이 수시로 변하는 인터넷 망이나 무선망과 같은 환경에 적절히 대응하지 못한다는 한계가 있다. 이러한 한계를 극복하기 위해 MPEG-4에서는 강화계층의 비트평면(bitplane) 부호화에 기반한 FGS(Fine Granularity Scalability)[1] 방식을 도입했다. 그러나 FGS 방식 및 전통적인 스케일러블 부호화 방식에서는 움직임 예측/보상 및 엔트로피 부호화 단계에서 이용할 수 있는 정보의 부족으로 인해 단일 계층으로 부호화하는 방식에 비해 성능이 떨어지는 단점이 꾸준히 제기되어 왔다. 이러한 문제점을 극복하기 위한 최근의 노력으로 스케일러블 특성 및 압축효율이 우수한 웨이블릿 변환을 이용한 영상 정보의 부호화 방식이 제안되었다.

본 논문에서는 웨이블릿을 이용한 고효율 스케일러블 비디오 부호화를 위한 시간적 필터링[2, 3] 방법을 제안한다. 본 논문의 구성은 다음과 같다. 2 절에서 움직임 예측/보상을 이용한 시간적 필터링 방법과 그에 따른 웨이블릿 변환 방법을 기술하고, 3 절에서 효율적

인 시간적 필터링을 위한 적응적인 가변블럭 크기 움직임 예측/보상 방법을 기술한다. 이어 4 절에서 실험결과를 보이고 5 절에서 결론을 맺는다.

II. Temporal Filtering 을 이용한 3 차원 웨이블릿 변환 부호화

3D 웨이블릿 부호화[2, 3]는 공간 영역에서의 웨이블릿 부호화를 시간영역으로 확장시킨 구조를 갖고 있다. 통상의 경우, 비디오 데이터는 많은 움직임을 포함하고 있고 이러한 영상에 바로 3 차원 변환을 수행하면 압축 효율이 크게 떨어지게 된다. 이런 단점을 해소하기 위해 제안된 방법이 움직임 예측/보상을 통한 3D 웨이블릿 부호화 방법이다.

전체적인 과정을 나타내는 블록 다이어그램은 다음과 같다(그림 1).

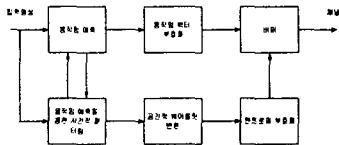


그림 1 움직임 예측/보상을 통한 3D 웨이블릿 변환 부호화 방법

먼저 입력된 영상은 움직임 예측을 통하여 얻어진 움직임 벡터가 가리키는 방향으로 시간적 필터링(MCTF; Motion Compensated Temporal Filtering)을 행함으로써 저역 통과 부대역과 고역 통과 부대역으로 분해하게 된다. 생성된 각각의 프레임들은 공간영역에서의 변환을 통해서 다시 공간적 부대역으로 분해가 되고 생성된 계수들은 양자화를 거친 후 엔트로피 부호화를 통해 압축된 비트스트림으로 부호화 된다. 시간적 필터링시에 사용되는 필터는 Karlsson 과 Vetterli 가 제안한 이후 현재까지 보편적으로 탭수가 작은 2-tap Haar 필터가 사용되고 있다[4]. 필터 탭수의 증가는 압축 효율을 높이는 데 도움을 줄 수 있지만 압축 이득이 미미할 뿐 아니라 복잡도의 증가와 높은 메모리 사용, 부호화 지연의 문제로 일반적으로 2-tap Haar 필터를 사용하고 있고 본 논문에서도 이 필터를 사용한다. 이 시간적 필터링의 과정을 그림 2 에 나타내었다. 그림 2 의 (a)는 현재 프레임과 이전 프레임간의 블록단위의 움직임을 통하여 얻어진 움직임 벡터를 나타내고 있다. 그림 2 의

(b)는 필터링을 거친 후 생성된 저역 통과 부대역과 고역 통과 부대역의 두 프레임을 나타낸다.

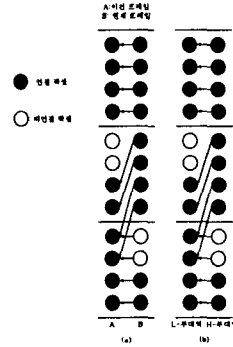


그림 2 움직임을 이용한 시간적 필터링 과정

그림처럼 움직임을 이용하게 되면 단순히 이전 프레임의 같은 위치의 픽셀과 필터링할 때에 비해서 신호의 에너지를 저역 통과 부대역으로 집중시킬 수 있다는 장점이 있다. 하지만 움직임 보상이 비연결 화소(unconnected pixel)가 생기게 되며 이는 압축 효율과 화질에 악 영향을 미치게 된다. 이런 비연결 화소의 경우에 필터링을 하지 못하기 때문에 식 (3), (4)와 같이 처리해 주고 있지만 비연결 화소의 수를 감소시켜야 압축 효율이 좀더 좋아진다. 지금까지의 2-tap Haar 필터를 이용한 필터링을 수식으로 표현하면 다음과 같다.

(i) For Connected Pixels:

$$L[m-d_m, n-d_n] = \frac{1}{\sqrt{2}} B[m, n] + \frac{1}{\sqrt{2}} A[m-d_m, n-d_n] \quad (1)$$

$$H[m, n] = \frac{1}{\sqrt{2}} B[m, n] - \frac{1}{\sqrt{2}} A[m-d_m, n-d_n] \quad (2)$$

(ii) For Unconnected Pixels:

$$L[m, n] = \frac{2A[m, n]}{\sqrt{2}} \quad (3)$$

$$H[m, n] = \frac{1}{\sqrt{2}} (B[m, n] - A[m-d_m, n-d_n]) \quad (4)$$

위의 수식에서 $L[m, n]$, $H[m, n]$ 은 필터링을 거친 이후의 저역 통과 부대역과 고역 통과 부대역 프레임, $A[m, n]$, $B[m, n]$ 는 이전 프레임과 현재 프레임, (d_m, d_n) 은 움직임 벡터를 나타낸다. 또한 기저함수가 정규 직교(orthonormal)한 성질을 갖도록 저대역 통과필

터와 고대역 통과필터는 각각 $[1/\sqrt{2}, 1/\sqrt{2}]$, $[1/\sqrt{2}, 1/\sqrt{2}]$ 이 된다.

본 논문에서는 움직임 예측과 보상과정을 발생 데이터량과 영상의 특성에 따라 가변적인 블록의 크기로 수행하는 방법을 제안한다. 그에 대한 내용은 다음절에서 기술한다.

III. 효과적인 Temporal Filtering을 위한 가변블록 움직임 예측 및 보상 방법

3D 웨이블릿 부호화기의 성능향상을 위해서는 움직임 예측과정에서 발생되는 비연결 화소의 수를 감소시킬 필요가 있다[2,3].

본 논문에서는 움직임 예측과 보상에 블록의 크기를 영상의 특성 및 전체적인 윌-왜곡을 고려해 가변적으로 적용함으로써 압축 효율을 높일 수 있는 방법을 제안한다. 일반적으로 비연결 화소의 수는 움직임 예측에 사용되는 블록의 크기가 커질수록 감소하고 작아지면 증가한다. 그러나 비연결 픽셀의 감소를 위해서 블록의 크기를 증가시키게 되면 영상의 세밀한 부분이나 움직임이 복잡한 경우 예측된 블록과의 에러가 커지게 되므로 결과적으로 이로 인한 부호화 손실이 생기게 된다. 또한 하나의 블록이 작은 크기의 여러 세부 블록으로 나뉘어 졌을 때 움직임 벡터의 데이터 량도 함께 증가하게 된다. 본 논문에서는 움직임 예측에 따른 블록 간의 에러와 움직임 벡터의 양을 고려하여 최적의 블록 크기를 결정함으로써 압축률을 높이는 방법을 기술한다.

비교적 단순한 영역이나 움직임이 작은 영역에서는 움직임 예측에 사용되는 블록의 크기가 커진다 해도 작은 블록을 사용했을 때에 비해서 에러의 증가가 크지 않고 적은 개수의 움직임 벡터로 영역을 표현할 수 있기 때문에 효율적이다. 반대로 영상의 세밀한 부분이나 움직임이 많은 영역에 대해서는 작은 크기의 블록을 사용하여 움직임 예측을 함으로써 예측 에러를 줄일 수 있기 때문에 영역의 특성에 따라 적절한 크기의 블록을 결정함으로써 압축률의 향상을 가져올 수 있다. 그렇기 때문에 본 논문에서는 16x16의 고정된 크기의 블록이 아닌 4x4에서 32x32까지의 가변적인 블록의 크기로 움직임 예측을 한다. 먼저 4x4 크기의 블록 단위로 움직임 예측을 행한 후, 현재 블록과 주변 3개의 블록의 움직임 벡터의 유사도를 측정한다. 다음의 식 (5)를 기준으로 주변 블록의 움직임 벡터가 유사한 값을 가지고

있다면 그 블록은 합하여 큰 블록으로 움직임 예측을 수행하고 이러한 과정을 반복하여 최적의 블록 크기와 해당 움직임 벡터를 구한다.

$$\min(D_{pred} + \lambda_{mv} R_{mv}) \quad (5)$$

이때, D_{pred} 는 움직임 보상 결과 발생하는 에러성분이고 식 (6), R_{mv} 는 해당하는 블록크기에 해당하는 움직임 벡터의 발생 비트수이다.

$$D_s(dx, dy) = \sum_{x,y} d(A[x, y] - B[x - dx, y - dy]) \quad (6)$$

이러한 과정을 거쳐 블록크기를 변화시켜 움직임 예측 및 보상을 수행하면 윌-왜곡 관점에서 최적의 블록크기와 움직임 벡터를 찾을 수 있고 이를 이용해 시간적 필터링을 수행하여 에너지를 저역 통과 부대에 집중시키고 비연결 화소의 수도 감소시킨다

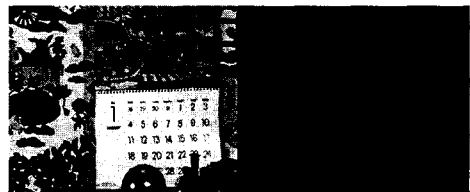
IV. 실험 결과

4.1 가변블록 움직임 예측 및 보상

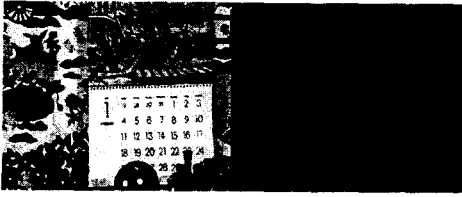
실험에 사용한 테스트 시퀀스는 Mobile & Calendar (CIF)와 Table-tennis(CIF)이고, 각각 30Hz, 15Hz의 프레임율에 대해 수행하였다. 표 1에 16x16 고정 크기의 블록 단위 움직임 예측 및 보상을 수행한 결과와 4x4에서 32x32까지의 가변 크기의 블록 단위로 움직임 예측 및 보상을 수행한 결과를 보였다. 결과는 움직임 예측 및 보상 후 1단계 시간적 필터링한 영상의 저역 통과 부대역 프레임의 화질과 비연결 화소수를 나타낸다. 그림 3은 결과 영상을 나타낸다.

표 1 고정블록을 사용한 경우와의 성능 비교

	PSNR (dB)	비연결 픽셀의 개수 (개)
고정블록 (16x16)	23.787	1565
가변블록	24.47	2548



(a) 고정블록을 사용하여 필터링을 거친 영상



(b) 가변 블록을 사용하여 필터링을 거친 영상

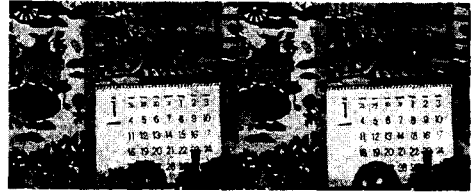


그림 6 고정 블록을 사용한 경우와 가변 블록을 사용한 경우의 복호된 영상 비교
(mobile & calendar, 384 kbps)

그림 3 MCTF후의 저역통과/고역통과 프레임

4.2 3D 웨이블릿 변환을 이용한 부호화

기존의 MPEG-4 단일 계층(single layer)으로 부호화된 결과와 제안한 방식을 비교하였다. 3D 웨이블릿 부호화는 Hsiang and Woods[2]의 결과를 이용하였고, 고정 크기의 움직임 예측 및 보상 결과와 가변 크기 움직임 예측 및 보상 결과를 비교하였다. 하나의 GOP는 16 프레임으로써 4단계의 시간적 필터링을 수행하였다.

그림 4 와 그림 5는 각각 목표 비트율에서의 화질을 그래프로 나타내고, 그림 6에서는 화질 비교를 위해 결과 프레임을 보였다.

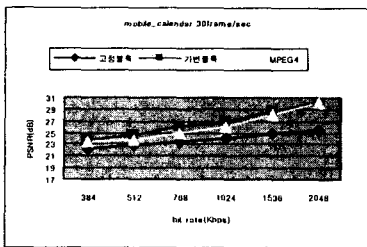


그림 4 고정블록을 사용한 경우와 MPEG4와의 PSNR 비교 (mobile & calendar)

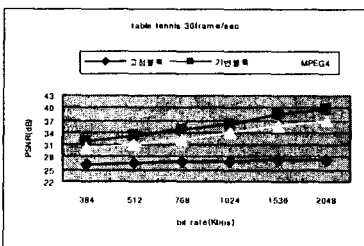


그림 5 고정블록을 사용한 경우와 MPEG-4와의 PSNR 비교 (table-tennis)

V. 결론

본 논문에서는 3D 웨이블릿 부호화기에서 움직임 예측시에 가변크기의 블록을 사용할 것을 제안하였다. 제안된 방법에서는 울-왜곡 관점에서 최적의 블록크기와 움직임 벡터를 찾음으로써 고정된 블록을 사용한 경우와 MPEG-4 를 사용하여 부호화를 행한 경우에 비해서 성능향상을 가져옴을 실험을 통해서 확인하였다.

참고문헌

- [1] W.Li, F. Ling, X. Chen, "Fine granularity scalability in MPEG-4 for streaming video," Proceedings of the IEEE International Symposium on Circuits and Systems, Vol.1, 2000, pp. 299-302.
- [2] S.-Ta Hsiang and J. W. Woods, "Embedded video coding using invertible motion compensated 3-D subband/wavelet filter bank," Signal Processing: Image Communication vol. 16, pp. 705-724, 2001.
- [3] S.-T. Hsiang, J.W. Woods, "Invertible three-dimensional analysis/synthesis system for video coding with half-pixel-accurate motion compensation," SPIE Conference on Visual Communications and Image Processing, San Jose, CA, Vol. 3653, 1999, pp.537-546
- [4] G.D. Karlsson, M. Vetterli, Three dimensional subband coding of video, in:International Conference on Acoustics, Speech, and Signal Processing, 1988, pp.1100-1103.