

다중해상도 움직임 보상과 블록 분할을 이용하는 웨이브렛 기반 동영상 부호화

양창모, 임태범, 이석필
전자부품연구원 디지털미디어연구센터
{cmyang, tlim, lspbio}@keti.re.kr

Wavelet based Video Coding using Multi-resolution Motion Compensation and Block Partition

Chang-Mo Yang, Tae-Beom Lim and Seok-Pil Lee
Digital Media Research Center, Korea Electronics Technology Institute (KETI)

요약

본 논문에서는 동영상을 효율적으로 부호화하기 위한 새로운 다중해상도 움직임 보상 방법과 잉여 양자화 방법을 제안한다. 본 논문에서 제안하는 동영상 부호화기는 다단계 이산웨이브렛 분해, 움직임 예측 및 움직임 보상, 블록 Tree의 구성 및 블록 분할, 적응적 산술 부호화기로 구성된다. 제안된 동영상 부호화기는 단순하면서도 낮은 연산량을 필요로 하며, 임베디드 특성과 SNR 계위 부호화 특성과 같은 좋은 기능을 제공한다. 또한, 기존에 제안되었던 이산웨이브렛변환을 이용하는 동영상 부호화 방법과 비교하여 우수한 성능을 제공한다.

I. 서론

H.26x 시리즈 권고안이나 MPEG 과 같은 기존의 표준에서는 영상을 일정한 크기를 가지는 블록으로 분할하여 이산여현변환(Discrete Cosine Transform)과 양자화를 수행하기 때문에 낮은 비트율에서 블록화 현상이 발생한다. 또한, 단일 압축 시스템에서 계위 부호화를 지원하기가 어렵다.

최근에는 이러한 문제점들을 해결하는 새로운 부호화 방법들이 연구되고 있다. 이러한 방법들 중 대표적인 것이 이산웨이브렛변환(Discrete Wavelet Transform)을 이용하는 부호화 방법이다.

이산웨이브렛변환을 이용하는 동영상 부호화기는 시간적 중복성을 줄이는 방법에 따라 크게 3-D 대역 부호화 방식과 예측 부호화 방식으로 분류된다.

3-D 대역 부호화 방식에서는 입력되는 동영상 화면 집합에 3-D 이산웨이브렛 분해를 수행한 후에, 기존의 정지영상 부호화 방법을 3-D 로 확장하여 동영상을 부호화한다[1,2].

비록 3-D 대역 부호화 방식이 기존의 이산웨이브렛변환을 이용하는 동영상 부호화 방법보다 나은 성능을 제공하지만, 움직임 객체 주변이 흐려지는 현상을 야기한다는 것과 많은 프레임 메모리를 필요로 한다는 단점을 가지고 있다.

움직임 보상을 이용하는 예측 부호화 방식은 움직임 예측 및 움직임 보상을 수행한 후, 잉여 양자화를 수행하는 방식이다. 대표적인 방법으로는 Lee 와 Oh 가 제안한 방법이 있다[3].

본 논문에서는 다중해상도 움직임 예측 및 움직임 보상(MRME/MRMC)[4]을 이용하는 새로운 예측 부호화 방법을 제안한다. 그림 1 은 본 논문에서 제안한 동영상 부호화기를 도시한 것이다. 먼저, 입력되는 동영상 화면에 2 차원 이산웨이브렛변환을 취한다. 그런 후에, 이산웨이브렛변환 영역에서 다중해상도 움직임 예측 및 움직임 보상을 수행한다. 이와 같이 변환 영역에서 움직임 예측 및 움직임 보상을 수행하게 되면 역변환을 가할 필요가 없으므로 연산 복잡도가 낮아진다. 움직임 탐색 방법으로는 전역 탐색(Full

Search)방법, 3 단계 탐색(Three-Step Search)방법 등이 사용될 수 있다. 본 논문에서는 연산량을 줄이기 위해 3 단계 탐색 방법을 사용하였다. 예측 오류는 이산웨이브렛변환 계수들 사이의 상관 관계와 블록분할 방법을 이용하여 양자화한다. 움직임 벡터와 양자화 값은 적응적 확률 모델을 이용하는 산술 부호기(Arithmetic Coder)를 이용하여 부호화 한다. 제안한 동영상 부호기에서의 부호화 비트율은 이산웨이브렛변환을 이용하는 양자화 방법의 임베디드(Embedded) 특성을 이용하여 제어한다.

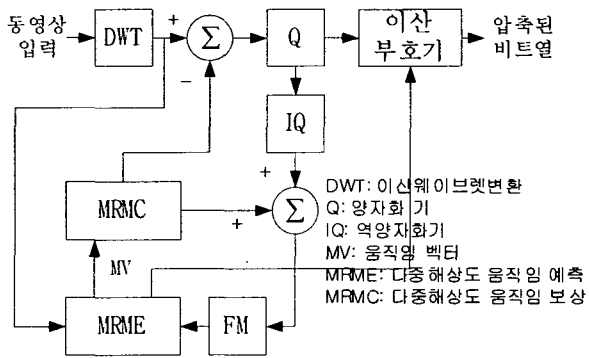


그림 1. 제안한 동영상 부호화기

II. 이산웨이브렛변환

이산웨이브렛변환은 입력되는 영상 신호에 다상 분해를 수행하는 변환 방식이며, 일반적으로 Quadrature Mirror Filter(QMF)를 이용하여 구현된다. 이산웨이브렛변환을 이용하여 영상의 수평 방향과 수직 방향으로 저주파 대역 필터와 고주파 대역 필터를 적용하고 2:1의 부표본화(Sub-sampling)를 수행하여 영상을 주파수 영역별로 4개의 대역으로 분할한다. 이러한 과정을 통해 얻어진 가장 낮은 주파수 대역에는 영상의 대부분의 신호 에너지가 몰려있으므로, 이 주파수 대역에, 앞서 설명한 분할 과정을 반복적으로 수행하여 다음 계층의 대역을 얻는다.

III. 움직임 예측 및 움직임 보상

동영상 부호화에 있어서 움직임은 매우 중요한 정보로 취급되고 있으며, 많은 동영상 부호화기에서 움직임 예측 및 움직임 보상을 채택하고 있다. 이는

움직임 정보를 통하여 동영상의 시간적 중복성을 제거함으로써, 보다 높은 압축효과를 기대할 수 있기 때문이다. 본 논문에서는 3 단계 움직임 탐색을 이용한 다중해상도 움직임 예측 및 움직임 보상 방법을 제안한다.

본 논문에서 제안한 다중해상도 움직임 예측 및 움직임 보상 방법을 그림 2에 도시하였다. 그림에서 보듯이, 최하위 주파수 대역은 다른 대역과 분할하여 독립적으로 반화소 움직임 예측 및 움직임 보상을 수행하였다. 이때 블록의 크기는 4×4 이고, 탐색 영역은 $[-4, +4]$ 이다. 최하위 주파수 대역을 제외한 나머지 대역에 대해서는 “부모-자식간 상관관계” [4,5,6]를 이용하여 가변 크기 블록 움직임 예측 및 움직임 보상을 수행하였다. 먼저, 최하위 주파수 대역과 같은 크기를 가지는 대역은 2×2 크기의 블록과 $[-2, +2]$ 의 탐색 영역을 이용하여 움직임 예측 및 움직임 보상을 수행한다. 그런 후에, 이 대역의 4배 크기를 가지는 대역에 대해서 “부모-자식간 상관관계”를 이용하여 움직임 벡터를 초기화한다. 이때 블록의 크기와 탐색 영역은 4배가된다. 또한 초기화된 움직임 벡터를 이용하여 탐색 영역이 새롭게 결정된다. 새롭게 결정된 탐색 영역을 이용하여 움직임 벡터를 정제하는 과정을 수행한다. 이러한 과정을 전 대역에 대해 반복적으로 적용하여, 움직임 예측과 움직임 보상을 수행한다.

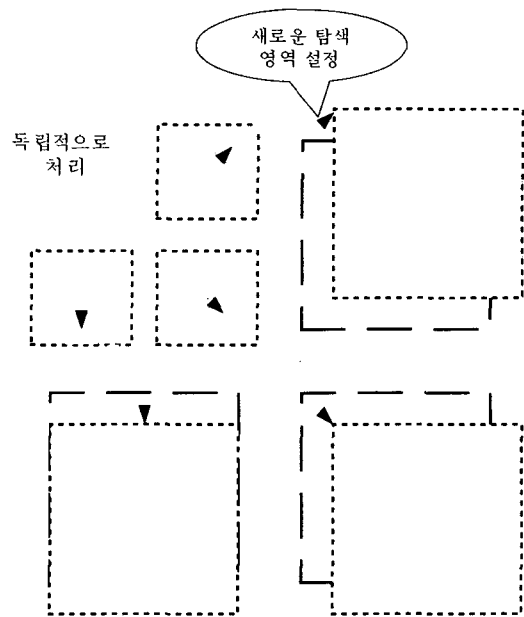


그림 2. 다중해상도 움직임 예측 및 움직임 보상

IV. 잉여 양자화

이산웨이브렛변환된 화면에 움직임 예측 및 움직임 보상을 끝내면, 예측 오류에 대해 잉여 양자화를 수행한다. 그런데, 각 대역에 위치한 예측 오류는 대부분 아주 작은 값을 가지게 되므로, 이러한 계수들은 0 으로 양자화해도 복원된 영상의 품질에 크게 영향을 미치지 않는다. 특히, 고주파 대역에서 비교적 큰 값을 가지는 계수의 발생 빈도는 매우 적으므로, 이러한 계수에 대해서만 부호화를 수행하고 그 위치정보를 수신측에 전송하는 것이 합리적이다. 각 대역간의 상관 관계를 이용하면, 이러한 중요 계수들의 위치 정보를 효율적으로 전송할 수 있다.

대역간 계수들의 상관 관계를 이용하는 대표적인 이산웨이브렛변환 기반 부호화기에는 EZW[6]와 SPIHT[5]이 있다. EZW 와 SPIHT 은 이산웨이브렛변환 계수의 “부모-자식간 상관관계”를 이용하여, 중요성 지도(Significance Map)를 만들고, 이를 부호화하는 방식을 사용한다. 중요성 지도는 지수적으로 감소하는 문턱치와의 비교를 통해 판단된 변환 계수의 중요성 여부를 나타낸 이진 비트평면(Bit Plane)이다.

본 논문에서는 이러한 종래의 기술들에 이용되었던 “부모-자식간 상관 관계”와 더불어 그림 3 과 같은 “사촌들간의 상관 관계”를 이용하여 영상을 부호화 한다. 먼저, 새로운 상관 관계를 이용하여 그림 4 의 왼쪽 그림과 같이 블록 Tree T 를 구성한다. 이렇게 구성된 Tree T 는 주어진 문턱치와 비교하여 Tree 내에 중요한 계수가 있다고 판단될 때에 그림 4 와 같이 보다 작은 크기를 가지는 블록 Tree $T_i (i \in 1, 2, 3, 4)$ 로 분할된다. 이러한 과정을 각 단계에서의 문턱치에 대해 더 이상 Tree 내에 중요한 계수가 없다고 판단될 때까지 반복하여 영상을 부호화 한다. 이때, Tree T 의 중요도는 다음과 같은 수식을 이용하여 판단한다.

$$S_n(T) = \begin{cases} 1, & \text{if } 2^n \leq \max_{(i,j) \in T} |c_{i,j}| < 2^{n+1} \\ 0, & \text{else} \end{cases} \quad (1)$$

위의 수식에서 T 는 Tree 를 의미하며 \mathcal{L} 은 주어지는 양자화 단계 n 에 대한 문턱치를 의미한다. 그리고 1 은 현재의 Tree T 가 중요한 계수를 포함하고 있다는 것을 의미하며, 0 은 중요한 계수를 포함하고 있지 않음을 의미한다. 유사한 방법을 이용하여, 좌표 (i, j) 세

위치하는 계수 $c_{i,j}$ 의 중요도는 다음과 같은 수식을 이용하여 판단한다.

$$S_n(c_{i,j}) = \begin{cases} 1, & \text{if } 2^n \leq |c_{i,j}| < 2^{n+1} \\ 0, & \text{else} \end{cases} \quad (2)$$



그림 3. 이산웨이브렛변환 계수의 사촌들간의 상관관계

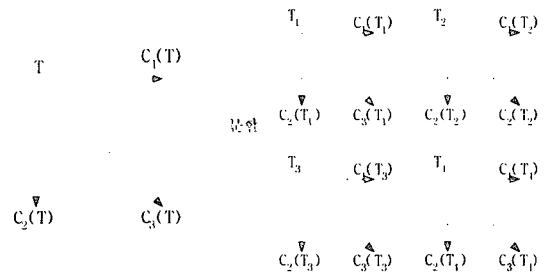


그림 4. 블록 Tree 의 구성과 블록 분할

V. 엔트로피 부호화

제안된 동영상 부호화에서의 결과 심벌은 적응적 산술 부호기(Adaptive Arithmetic Coder)를 이용하여 부호화 한다. 계수의 중요도 부호화를 위해서는 동쪽, 서쪽, 남쪽, 북쪽에 위치한 4 개의 인접계수와 하나의 부모 계수의 중요도 값을 이용하였고, 계수의 부호 부호화를 위해서는 동쪽, 서쪽, 남쪽, 북쪽에 위치한 4 개의 인접 계수들의 부호 정보를 이용한다. 계수의 중요도 부호화에 부모 계수의 중요도 값이 사용된 것은 “부모-자식간 상관 관계”를 이용하기 위함이다. Tree 의 중요도, 정제단계에서의 결과 심벌, 그리고 움직임 벡터를 부호화하기 위해서는 고정된 하나의 확률모델을 이용한다.

VI. 실험 결과 및 분석

본 논문에서의 이산웨이브렛변환에는 9/7-tap bi-orthogonal 선형위상 필터가 사용되었다. 복원된 영상 신호의 성능을 측정하는 방법으로는 PSNR(Peak Signal-to-Noise Ratio)을 이용하였다.

표 1 은 제안한 동영상 부호화기의 성능을 10 fps(frame per second), 30 kbps, 60 kbps 의 실험조건에서 3-D SPIHT[2]과 비교한 것을 보여준다. 성능평가를 위해서 MPEG 표준에서 이용되는 QCIF 크기의 HALL MONITOR 와 MOTHER AND DAUGHTER 동영상 시퀀스를 사용하였다. 각각의 동영상 시퀀스는 30 fps, 4:2:0 형식으로 구성되어 있다. 원 동영상 시퀀스가 30 fps 로 구성되어 있으므로, 세 화면마다 하나의 화면을 선택하여 실험하였다. 표 1 에서 알 수 있듯이, 제안한 동영상 부호화기는 3-D SPIHT 과 비교하여, HALL MONITOR 동영상 시퀀스에 대해서는 비슷하거나 다소 좋지 못한 성능을 보이지만 MOTHER AND DAUGHTER 와 같이 움직임이 복잡한 동영상 시퀀스에 대해서는 많은 부호화 이득을 얻을 수 있었다.

표 1. 평균 PSNR 비교

부호화 방법	Bit Rate	MOTHER AND DAUGHTER	HALL MONITOR
제안한 방법	30 kbps	32.9 dB	32.9 dB
	60 kbps	35.8 dB	37.3 dB
3-D SPIHT	30 kbps	32.7 dB	32.9 dB
	60 kbps	35.5 dB	37.9 dB
MC 3-D SPIHT	30 kbps	32.7 dB	32.3 dB
	60 kbps	35.6 dB	37.3 dB

VII. 결론

본 논문에서는 블록화 현상도 발생하지 않고 단일 압축 시스템에서 계위 부호화를 지원하기에도 용이한 이산웨이브렛변환 기반 동영상 부호화기를 제안하였다. 제안된 동영상 부호화기는 크게 이산웨이브렛변환, 움직임 예측 및 움직임 보상, 잉여 양자화, 엔트로피 부호화의 4 가지로 나눌 수 있다. 이산 웨이브렛 변환에는 구현이 간단하면서도 효율적인 QMF 방식을 이용하였다. 움직임 예측 및 움직임 보상은 다중해상도 움직임 예측 및 움직임 보상 방법이

이용되었다. 잉여 양자화 방법으로는 이산웨이브렛 계수들간에 존재하는 상관 관계와 블록 기반 Tree 분할을 이용하는 양자화 방법을 사용하였다. 엔트로피 부호화에는 연산이 빠르면서도 적응적으로 확률을 모델링할 수 있는 산술 부호화기를 사용하였다. 실험을 통해 보인 것처럼, 본 논문에서 제안한 동영상 압축 부호화기를 이용하여 기존의 3-D SPIHT 에 비해 향상된 압축 효율을 얻을 수 있었으며, 특히 움직임이 많은 동영상에서 많은 부호화 이득을 얻을 수 있었다. 또한 본 논문에서 제안한 동영상 부호화기는 3-D SPIHT 과 비교하여 메모리 사용량이 적다는 장점을 가지고 있다.

참고문헌

- [1] J. Y. Tham, S. Ranganath and A. A. Kassim, "Highly Scalable Wavelet-based Video Codec for Very Low Bit-Rate Environment," *IEEE Journal on Selected Areas in Comm.*, vol. 16, no. 1, pp. 12-27, Jan. 1998.
- [2] B. J. Kim, Z. Xiang and W. A. Pearlman, "Low Bit-Rate Scalable Video Coding with 3-D Set Partitioning in Hierarchical Trees," *IEEE Trans. on Circuits and Systems for Video Tech.*, vol. 10, no. 8, pp. 1374-1387, Dec. 2000.
- [3] J. Y. Lee, H. S. Oh and S. J. Ko, "Motion-Compensated Layered Video Coding for Playback Scalability," *IEEE Trans. on Circuits and Systems for Video Tech.*, vol. 11, no. 5, pp. 619-628, May 2001.
- [4] S. Zafar, Y. Zhang, and B. Jabbari, "Multi-scale Video Representation using Multi-resolution Motion Compensation and Wavelet Decomposition," *IEEE Journal on Selected Areas in Comm.*, vol. 11, no. 1, pp. 24-35, Jan. 1993.
- [5] A. Said and W. Pearlman, "A New, Fast and Efficient Image Codec based on Set Partitioning in Hierarchical Trees," *IEEE Trans. on Circuits and Systems for Video Tech.*, vol. 6, no. 3, pp. 243-250, June 1996.
- [6] J. Shapiro, "Embedded Image Coding using Zerotrees of Wavelet Coefficients," *IEEE Trans. on Signal Proc.*, vol. 41, no. 12, pp. 3445-3462, Dec. 1993.