

변형된 IDCT 기저 함수를 이용한 압축된 동영상의 하향 전환기법

Down Conversion Algorithm for Compressed Video Sequence Using a Modified IDCT Basis Function in Transform Domain

김 명 준, 송 병 철, 장 성 규, 나 종 범

한국과학기술원 전기및전자공학과 영상시스템연구소

Myung Jun Kim, Byung Cheol Song, Sung Kyu Jang, and Jong Beom Ra

Dept. of Electrical Engineering, KAIST

요 약

본 논문은 DCT (Discrete Cosine Transform) 영역에서의 압축 동영상 하향 전환기법 (down conversion)을 제안한다. DCT 영역에서의 하향 전환이 완전 복호화한 후 공간 영역에서 하향 전환하는 것보다 계산량 측면에서 상당한 이점이 있다. 또한 복호기 루프 내에서 영상 크기가 줄기 때문에 메모리의 부담도 덜 수 있다. 가장 간단한 방법으로 복원된 영상의 화질이 약간 떨어지더라도 계산량과 메모리를 줄이기 위해 8x8 DCT 블록의 저주파 영역의 4x4 DCT 계수만을 추출하여 4x4 IDCT 하는 기법이 널리 알려져 있다. 본 논문에서는 변형된 4x4 IDCT 기저 함수를 이용한 새로운 DCT 영역에서의 하향 전환 기법을 제안한다. 모의 실험을 통해 제안한 기법이 기존의 DCT 영역에서의 하향 전환기법과 같은 계산량 및 메모리로 향상된 PSNR을 갖는다는 것을 보인다.

I. 서 론

MPEG 과 H.261 같은 동영상 부호화 표준에서 부호화 효율을 높이기 위해 DCT가 이용된다 [1-2]. VLSI 기술이 급속히 발달하고 멀티미디어 응용 분야들이 점점 늘어남에 따라 실시간으로 동영상을 축소하거나 움직임을 추정하는 기술 등이 요구되고 있다. 그러나 방대한 동영상 데이터의 실시간 처리를 위해서는 두 가지 해결할 문제가 있다. 하나는 동영상 압축/복원을 위한 많은 계산량이며, 다른 하나는 처리 데이터의 높은 비트율이다. 따라서 복호화를 거쳐 공간 영역에서 영상 처리를 하는 방식으로는 실시간 처리가 힘들다. 이러한 이유 때문에 부분적인 복호화만을 수행하여 DCT 영역에서 모든 처리를 함으로써 최소한의 계산량만 갖는 고속 알고리즘에 관한 연구가 최근 활발히 진행되고 있다.

최근 들어 동영상의 축소 전환을 변환 영역에서 수행하는 여러 기법들이 개발되었다 [3]. 변환 영역에서의 축소 (TDD; transform domain decimation)란 DCT 영역에서 고 해상도를 갖는 영상의 비트열로부터 저 해상도의 영상을 복원해내는 기법을 말한다. TDD는 그림 1의 공간 영역에서의 축소 (SDD; spatial domain decimation)보다 적은 계산량을 갖는다 [1]. 이는 TDD의 모든 처리 과정이 DCT 영역에서 이루어 지기 때문이다.

그림 2에 나타나 있는 4x4 cut은 가장 간단한 TDD 방법으로서 8x8 DCT 블록으로부터 저 대역 4x4 계수들만 추출하여 4x4 IDCT (Inverse DCT)를 수행한다. 이 방법은 최소의 메모리와 계산량을 갖는다는 장점이 있지만 복원 영상

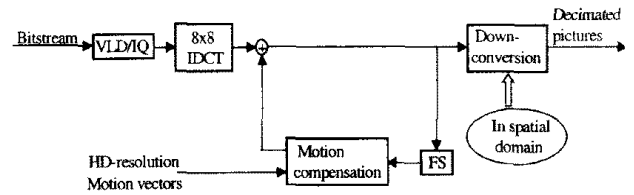


그림 1: 공간 영역에서의 축소 (SDD)의 블록도.

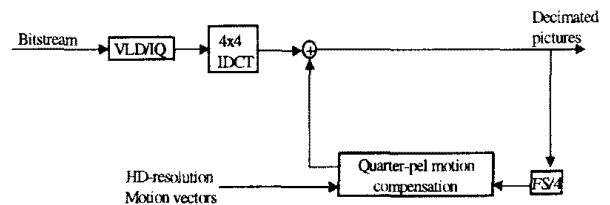


그림 2: 4x4 cut의 블록도.

의 화질이 상당히 떨어지는 단점이 있다.

최근 4x4 cut의 개념을 16x16 블록단위로 확대한 주파수 합성 (frequency synthesis)이라는 새로운 기법이 개발된 바 있다 [4-5]. 주파수 합성 기법은 인접한 네 개의 8x8 DCT 블록으로부터 합성된 16x16 DCT 블록의 왼쪽 상위 8x8 DCT 계수들에 대해 8x8 IDCT를 수행함으로써 축소된 영상을 얻는 방법이다. 이 방법이 4x4 cut보다 신뢰할만한 성능을 보이지만 16x16 DCT 블록을 합성을 위해 필요한 계산량은 일반적인 복호기에 상당한 부담이 된다.

DCT 계수의 상대적 중요도를 고려한 TDD도 있다 [6]. 이 방법은 DCT 계수의 절대값을 해당 계수의 중요도로 간주하고 중요도가 큰 순으로 계수를 추출한다. 그러나 특정 고주파 성분을 무시함으로써 축소 영상의 화질이 크게 저하될 수도 있는 단점을 지닌다.

본 논문은 일반적인 4x4 IDCT 기저함수 (basis function)에서 변형된 4x4 IDCT 기저함수를 갖는 새로운 TDD를 제안한다. 제안한 기저함수는 DCT 영역에서의 조날 마스크 (zonal mask)에 의한 저 대역 필터링 (low pass filtering) [7]과 MPEG에서 제시된 하향 전환 (down conversion) 필터링 효과를 가지는 서브 샘플링 (sub-sampling)에 의해 얻어진다 [1]. 제안한 하향 전환 기법은 기존의 4x4 cut과 동일한 메모리와 계산량을 가지고 더 나은 화질을 제공한다. 2장에서는 제안한 TDD를 이용한 동영상 복호기를 보인다. 3장에서는 실험 결과를 통해 기존의 방법과 비교를 통해 성능을 평가하며 4장에서 결론을 맺는다.

*본 연구는 산자부의 연구비 지원으로 이루어짐.

II. 제안한 알고리즘

II.1. DCT 계수 4x4 에 대한 필터링

일반적인 MPEG-2 복호기처럼 제안한 복호기에서의 모든 처리는 8x8 블록 단위로 한다. I 프레임 (intra frame)의 경우 각 8x8 블록은 제안한 4x4 IDCT 기법에 의해 간단히 축소 전환된다. B 나 P 프레임의 경우는 저 해상도의 참조 프레임 (reference frame)으로부터 움직임 보상된 4x4 블록과 제안한 4x4 IDCT 커널을 이용해 축소된 오류 블록을 더함으로써 복원된 4x4 블록을 얻는다. 움직임 보상된 블록을 얻기 위해 고 해상도의 움직임 벡터가 그대로 이용된다. 따라서 1/4 화소 단위 움직임 보상 (quarter-pel motion compensation)을 도입한다 [8]. 그림 3은 제안한 복호기이다.

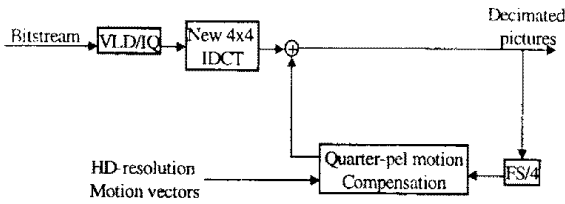


그림 3: 변형된 4x4 IDCT 기저함수를 이용한 MPEG-2 하향 전환 복호기의 구조

변형된 4x4 IDCT 기저함수는 DCT 영역에서의 조날 마스크에 의한 저 대역 필터링과 MPEG이 제시한 하향 전환 (down conversion) 필터의 특성을 가지는 서브 샘플링에 의해 얻어진다. $\{v(k,l) | 0 \leq k,l \leq 7\}$ 의 입력 8x8 DCT 블록에 조날 마스크와 8x8 IDCT를 적용하면 다음과 같은 결과가 얻어진다.

$$u(m,n) = \sum_{k=0}^3 \sum_{l=0}^3 v(k,l) C_8^{k(2m+1)} C_8^{l(2n+1)}, \quad 0 \leq m,n \leq 7 \quad (1)$$

여기에서 $C_8^{k(2m+1)}$ 는 8x8 IDCT 기저함수이며 다음과 같이 정의된다.

$$C_8^{k(2m+1)} = \alpha(k) \cos \frac{\pi(2m+1)k}{16} \quad (2)$$

여기에서 $\alpha(0) = \sqrt{\frac{1}{8}}$, $\alpha(k) = \sqrt{\frac{1}{4}}$, $1 \leq k \leq 7$ 이다.

식 (1)의 8x8 공간 영역 블록에 대해 MPEG이 제시한 하향 전환 필터를 적용하면 축소된 영상이 구해진다 (그림 4) [1]. 그림 4에서 행렬 A는 DCT 유형 (프레임 DCT, top 필드 DCT, bottom 필드 DCT)에 따라 달라지며, 행렬 B는 DCT 유형에 관계없이 일정하다. 이는 MPEG-2의 필드 특성에 기인한다.

그런데 8x8 블록 단위로 수행되어짐에 의해 그림 5과 같이 블록의 경계 부근에서 필터의 입력 데이터가 존재하지 않는 문제가 발생하는데, 이런 경우에는 주위의 화소값을 이용하였다 [8].

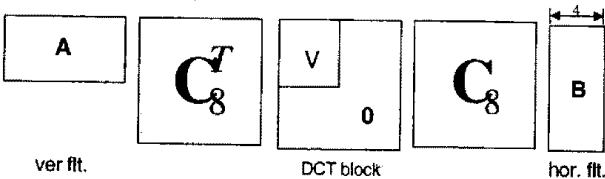


그림 4: DCT 계수 4x4 에 대한 필터링

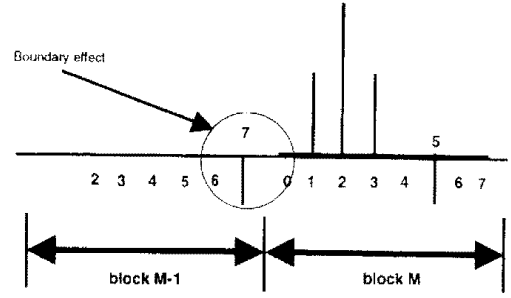


그림 5: 인접 블록의 경계에서의 하향 전환 필터링.

II.2. 수학적 분석 및 계산량 감소 방안

앞 절에서는 4x4 DCT 블록에 대한 두 단계 필터링, 즉 조날 마스크에 의한 저 대역 필터링과 하향 전환 필터링 효과의 서브 샘플링에 대해서 살펴 보았다. 이 절에서는 이러한 수행 과정을 수학적으로 분석하고 계산량의 감소 방안을 제안한다. 4x4 축소 블록 $\{\tilde{u}(m',n') | 0 \leq m',n' \leq 3\}$ 을 두 단계 필터링에 의해 얻을 수 있는데 분석을 쉽게 하기 위해 DCT 블록의 유형이 top 필드 DCT인 경우 수직 방향에 대한 처리만을 살펴보자.

i) $m' = 0$ 인 경우

$$\tilde{u}(m') = \frac{1}{256} \sum_{k=0}^3 v(k) \alpha(k) \begin{bmatrix} 138 * \cos\left(\frac{\pi(4m'+1)k}{16}\right) \\ +88 * \left\{ \cos\left(\frac{\pi(4m'+1)k}{16}\right) + \cos\left(\frac{\pi(4m'+3)k}{16}\right) \right\} \\ -29 * \left\{ \cos\left(\frac{\pi(4m'+1)k}{16}\right) + \cos\left(\frac{\pi(4m'+7)k}{16}\right) \right\} \end{bmatrix} = \sum_{k=0}^3 v(k) \alpha'(m',k) \cos\left(\frac{\pi(4m'+1)k}{16}\right), \quad 0 \leq k \leq 3$$

ii) $m' = 1$ 인 경우

$$\tilde{u}(m') = \frac{1}{256} \sum_{k=0}^3 v(k) \alpha(k) \begin{bmatrix} 138 * \cos\left(\frac{\pi(4m'+1)k}{16}\right) \\ +88 * \left\{ \cos\left(\frac{\pi(4m'-1)k}{16}\right) + \cos\left(\frac{\pi(4m'+3)k}{16}\right) \right\} \\ -29 * \left\{ \cos\left(\frac{\pi(4m'-3)k}{16}\right) + \cos\left(\frac{\pi(4m'+7)k}{16}\right) \right\} \end{bmatrix} = \sum_{k=0}^3 v(k) \alpha'(m',k) \cos\left(\frac{\pi(4m'+1)k}{16}\right), \quad 0 \leq k \leq 3$$

iii) $m' = 2$ 인 경우

$$\tilde{u}(m') = \frac{1}{256} \sum_{k=0}^3 v(k) \alpha(k) \begin{bmatrix} 138 * \cos\left(\frac{\pi(4m'+1)k}{16}\right) \\ +88 * \left\{ \cos\left(\frac{\pi(4m'-1)k}{16}\right) + \cos\left(\frac{\pi(4m'+3)k}{16}\right) \right\} \\ -29 * \left\{ \cos\left(\frac{\pi(4m'-5)k}{16}\right) + \cos\left(\frac{\pi(4m'+7)k}{16}\right) \right\} \end{bmatrix} = \sum_{k=0}^3 v(k) \alpha'(m',k) \cos\left(\frac{\pi(4m'+1)k}{16}\right), \quad 0 \leq k \leq 3$$

iv) $m' = 3$ 인 경우

$$\begin{aligned} \tilde{u}(m') &= \frac{1}{256} \sum_{k=0}^3 v(k) \alpha(k) + 88 * \left\{ \cos\left(\frac{\pi(4m'-1)}{16}k\right) + \cos\left(\frac{\pi(4m'+3)}{16}k\right) \right\} \\ &\quad - 29 * \left\{ \cos\left(\frac{\pi(4m'-5)}{16}k\right) + \cos\left(\frac{\pi(4m'+3)}{16}k\right) \right\} \\ &= \sum_{k=0}^3 v(k) \alpha'(m',k) \cos\left(\frac{\pi(4m'+1)}{16}k\right), 0 \leq k \leq 3 \end{aligned} \quad (3)$$

식 (3)의 $\alpha'(m',k)$ 는 식 (2)의 $\alpha(k)$ 와 달리 m' 의 값에 따라 변하며 그 값은 다음과 같다.

$$\alpha'(m',k) = \begin{bmatrix} 1 & 1.0382 & 1.0245 & 0.7638 \\ 1 & 1.0871 & 1.1843 & 1.0101 \\ 1 & 1.0871 & 1.1843 & 1.0101 \\ 1 & 1.0669 & 1.0245 & 1.6047 \end{bmatrix} \quad (4)$$

이 값들이 4x4 DCT 블록의 원소들에 가중치를 주게 되어 보다 나은 저 대역 필터링에 의한 성능 향상을 얻을 수 있다. 식 (3)을 그림 6과 같이 나타낼 수 있다.

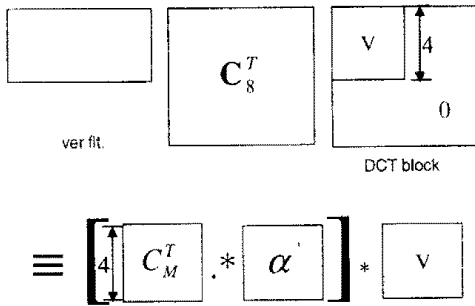


그림 6: 새로운 4x4 IDCT 커널.

여기에서 $\{C_M^T(m',k) | 0 \leq m', k \leq 3\}$ 은 다음과 같다.

$$C_M^T(m',k) = \cos\left(\frac{\pi(4m'+1)}{16}k\right) \quad (5)$$

8x8 블록의 DCT 유형이 프레임 DCT나 bottom 필드 DCT인 경우도 비슷한 방법을 적용할 수 있다. 그런데 MPEG-2의 필드 특성에 따라 DCT 유형별로 그림 6에서 보이는 것처럼 위치 보상이 필요하다. bottom 필드의 경우 반 화소만큼 이동된 위치의 화소 값을 취해야 하며, 이로 인해 $\{C_M^T(m',k) | 0 \leq m', k \leq 3\}$ 는 식 (6)과 같이 표현된다.

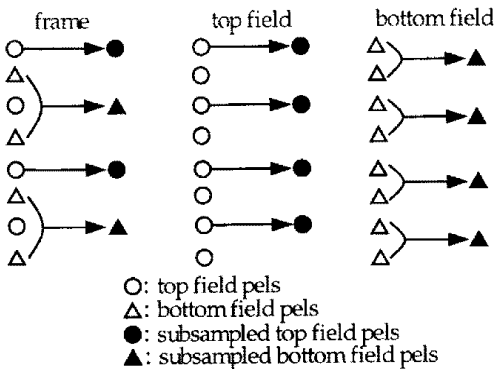


그림 7: DCT 유형에 따른 위치 보상.

$$C_M^T(m',k) = \cos\left(\frac{\pi(4m'+2)}{16}k\right) \quad (6)$$

한편 $\{\alpha'(m',k) | 0 \leq m', k \leq 3\}$ 는 k 와 m' 의 함수이므로 4x4 IDCT 보다 계산량이 증가하게 된다. 따라서 $\alpha'(m',k)$ 가 k 에만 의존하며, DCT 유형 별로 최소 오차를 갖도록 $\{\alpha''(k) | 0 \leq k \leq 3\}$ 를 구했다. 결과식이 (7)에 나타나 있다.

$$\begin{aligned} \alpha''(k) &= [1 \quad 0.9617 \quad 0.6540 \quad 0.1968] \quad \text{for frame} \\ &= [1 \quad 1.0565 \quad 1.1044 \quad 0.9362] \quad \text{for top field} \\ &= [1 \quad 0.9652 \quad 0.8623 \quad 0.6983] \quad \text{for bottom field} \end{aligned} \quad (7)$$

식 (5)와 (6)의 IDCT 기저함수는 Lee's IDCT [9]에 의해 빠른 알고리즘으로 구현될 수 있음을 쉽게 증명할 수 있다. 그리고 수평 방향의 IDCT 기저함수는 top 필드의 수직 방향의 IDCT 기저함수의 transpose 이다.

지금까지 식 (5), (6), (7)의 4-point IDCT 기저함수를 이용한 새로운 TDD 방법을 제안하였다. 본 논문에서는 앞으로 제안한 방법을 개선된 4x4 cut (modified 4x4 cut)이라고 부른다. 1/4 정확도의 움직임 보상을 사용할 경우 기존 4x4 cut에 의한 화소 위치는 그것의 원 위치와 잘 일치하지 않는 반면, 개선된 4x4 cut을 사용하여 얻어진 화소는 보다 정확히 일치한다. 이는 기존의 4x4 cut과는 달리 개선된 4x4 cut은 해당 위치의 화소를 직접 서브샘플링(sub-sampling)하는 효과를 지니기 때문이다. 또한 제안된 방법은 기존의 4x4 cut과 같은 계산량과 메모리만을 필요로 한다.

III. 실험결과

실험을 위해 "car", "football", "cheerleaders", "susie"의 4개의 CCIR 601 포맷 영상 (720x480; 100 frame)과 "basketball", "tennis"의 HDTV 영상 (1920x1080; 15 프레임)을 이용했다. MPEG-2 부호기에서 비트율을 CCIR 601 영상에 대해서는 4Mbps, HDTV 영상에 대해서는 18Mbps로 했다. GOP (group of picture) 크기는 15 프레임이고 P 프레임 간의 거리는 3 프레임이다. 완전 복호화 후 영상을 축소하는 SDD 방법이 가장 좋은 화질과 PSNR을 가진다고 가정하고 기존의 4x4 cut과 제안된 4x4 cut의 결과를 비교했다. SDD에서는 MPEG 2의 이미지 축소 필터 (decimation filter)를 사용했다 [1]. 표 1은 평균 PSNR에 의한 성능 비교 결과를 보여준다. 표 1에서 개선된 4x4 cut이 기존 4x4 cut보다 1dB 이상의 PSNR 향상을 가져옴을 알 수 있다. 그림 7와 8은 PSNR 그래프도 이러한 사실을 뒷받침한다.

IV. 결론

본 논문은 DCT 영역에서의 압축 영상의 축소를 위해 개선된 4x4 cut이라는 새로운 방법을 제안했다. 새로운 4x4 IDCT 기저함수는 DCT 영역에서 조날 마스크와 MPEG이 제시한 하향 전환 (down conversion) 필터의 합성에 의해 얻어진다. 모의 실험 결과로부터 제안한 개선된 4x4 cut이 기존의 4x4 cut과 동일한 복잡도와 메모리를 유지하면서 1dB 이상의 성능 향상을 보임을 증명하였다.

참고문헌

1. MPEG-2 IS, Video Recommendation ITU-T H.262, ISO/IEC 13818-2, Jan. 20, 1995.
2. CCITT Study group XV, "Draft revision of recommendation H.261-Video codec for audio visual services at px64 kbps," Temporary Document 5-E, July 1990.
3. S. Ng, "Low resolution HDTV receivers," *US Patent 5,262,854*, Nov. 1993.
4. A. Vetro, J. Bao, H. Sun, and T. Poon, "Frequency domain down conversion of HDTV using adaptive motion compensation," *IEEE Proc. ICIP*, Santa Barbara, CA, 1997.
5. A. Vetro and H. Sun, "Generalized motion compensation for drift reduction," *SPIE Proc. VCIP*, vol. 3309, pp. 484-495, 1998.
6. K. K. Pang, H. G. Lim, S. Dunstan, and J. M. Badcock, "Frequency domain decimation and interpolation techniques," *PCS '96*, Melbourne, Australia, March 1996.
7. A. K. Jain, *Fundamentals of digital image processing*, Englewood Cliffs, NJ: Prentice-Hall, Chapt. 7, pp. 256, 1989.
8. H. Sun, "Hierarchical decoder for MPEG compressed video data," *IEEE Trans. On Consumer Electronics*, Aug. 1993.

표 1: 제안한 방법과 기존의 4x4 cut 의 평균 PSNR [dB] 비교

		Conventional 4x4 cut	Modified 4x4 cut
CCIR 601 SEQUENCES	FOOT	26.25	27.31
	CAR	27.11	28.37
	CHEER	24.78	26.08
	SUSIE	34.15	35.65
HDTV SEQUENCES	BASKET	25.14	26.95
	TENNIS	24.62	25.64

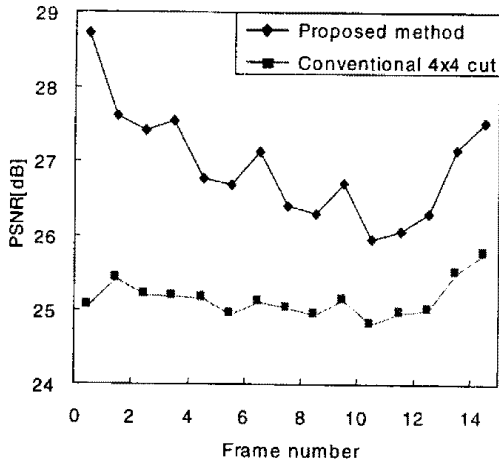


그림 8: Basket 영상에 대한 실험결과.

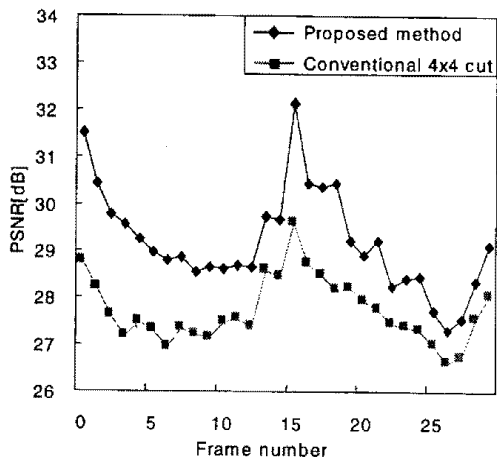


그림 9: Car 영상에 대한 실험결과.