

# 압축된 동영상에서 AC 예측 기법을 이용한 DC 영상 추출 기법

김 성 득, 나 종 범

한국과학기술원 전기 및 전자공학과

Tel: 042-869-5434, Fax: 042-869-8360

E-mail: sdkim@issserver.kaist.ac.kr, jbra@ee.kaist.ac.kr

## A DC IMAGE EXTRACTION SCHEME USING AC PREDICTION IN COMPRESSED VIDEO SEQUENCES

Sung Deuk Kim and Jong Beom Ra

Department of Electrical Engineering, KAIST

Tel: 042-869-5434, Fax: 042-869-8360

E-mail: sdkim@issserver.kaist.ac.kr, jbra@ee.kaist.ac.kr

### ABSTRACT

Video data is usually stored in a compressed format in order to reduce the storage space. For efficient browsing, searching, and retrieval of compressed video sequences, size-reduced images (or DC images which are formed with block DC coefficients) are generally preferred to avoid unnecessary computational complexity. In this paper, we propose a DC image extraction scheme appropriate for scene analysis and efficient browsing of compressed video sequences. The proposed algorithm utilizes predicted low frequency AC coefficients to achieve better approximation and to reduce the error drift. Due to the AC prediction based on a quadratic surface model, the proposed scheme requires no additional memory compared with the previous zero-order or first-order approximation scheme. Simulation results show that the proposed scheme achieves better subjective and objective quality with minor additional operations.

### 1. 서론

비디오 데이터의 효율적인 보관 및 전송을 위해 MPEG과 같은 동영상 압축에 관한 국제 표준안이 널리 사용되고 있다. MPEG은 시각적으로 좋은 화질을 유지하면서 비디오 자료를 표현하는데 필요한 비트량을 혁신적으로 감소시키는 압축 기법이지만, 압축된 자료는 영상처리의 목적으로는 다루기가 손쉽지 않다. 특히, 압축된 비디오 자료의 수가 날로 급증함에 따라 빠른 검색 및 복원을 위한 효과적인 알고리즘에 대해 많은 연구들이 진행되고 있다. [1,2]

실제 복원 영상보다 작은 크기의 영상 혹은 블록의 DC 계수 값만으로 이루어진 DC 영상은 영상을 분석하고 대략적으로 표현하는데 필요한 전반적인 요소들을 지니고 있는 것으로 알려져 있다 [2,3]. 더욱이 DC 영상의 추출은 완전한 크기로의 복원에 비해 보다 적은 메모리와 연산능력만으로 수행이 가능하므로 MPEG

기반으로 압축된 동영상의 효과적인 탐색에 사용되고 있다.

I-프레임에 대한 DC 영상 추출은 간단하다. 하지만 P-프레임 및 B-프레임의 경우, DC 영상을 얻기 위해서는 움직임 정보가 사용되어야 한다. 정확한 DC 영상을 얻기 위해서는 적지 않은 메모리와 연산능력을 필요로 한다. 이런 문제를 극복하기 위해 zero-order 및 first-order 근사 기법이 사용되었다 [4]. 하지만 이러한 방법들은 근사화 과정에서 AC 성분을 고려하지 않기 때문에 프레임들이 처리될수록 추출된 DC 영상의 화질이 점차 나빠지게 된다.

본 논문에서는 AC 계수의 예측기법을 이용한 DC 영상 추출기법을 제안한다. 인접한 블록의 DC 성분들로부터 AC 계수들을 예측하기 위하여 영상을 이차 평면 모델로 가정한다. 이 모델로부터 5개의 낮은 주파수 AC 계수들을 예측하고, 예측된 AC 계수들이 DC 영상 복원에 사용된다. 여기서 AC 계수들은 실제적으로는 이전 프레임의 복원된 DC 영상으로부터 얻어지기 때문에 제안하는 방법에서는 AC 성분을 위한 추가적인 메모리를 필요로 하지 않는다. 2장에서는 정확한 DC 영상을 얻기 위한 방법, zero-order 및 first-order 근사 기법들을 간략히 살펴본다. 3장에서는 제안하는 DC 영상 추출 기법을 자세히 설명한다. 4장에서 모의 실험 결과를 보인 후 5장에서 결론을 맺는다.

### 2. 기존의 DC 영상 추출 방법들

이 장에서는 정확한 DC 영상 추출기법을 먼저 소개하고 그 후 zero-order 및 first-order 근사기법을 간략히 소개한다.

#### 2.1. 정확한 DC 영상 추출 기법

I-프레임으로 압축된 프레임에서의 DC 영상 추출은 손쉽게 이루어진다. 하지만 P- 및 B- 프레임에서의 DC 영상을 얻어내기 위해서는 그림 1과 같이 움직임 정보가

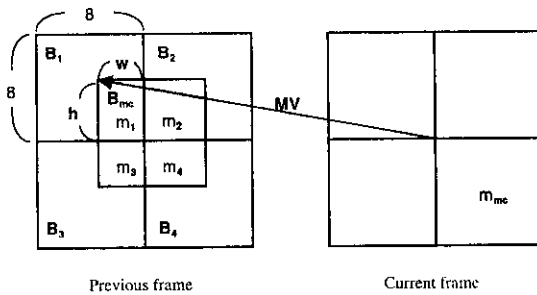


그림 1. P-프레임에서의 움직임 보상

필요하다. 움직임 보상된 블록의 DC 값은 공간영역에서 다음과 같이 얻어진다.

$$m_{mc} = \frac{1}{64} \left[ \begin{matrix} h \cdot w \cdot m_1 + h \cdot (8-w) \cdot m_2 \\ + (8-h) \cdot w \cdot m_3 + (8-h) \cdot (8-w) \cdot m_4 \end{matrix} \right] \quad (1)$$

여기서

$$\begin{aligned} m_1 &= \text{MEAN}(B_1 \cap B_{mc}), m_2 = \text{MEAN}(B_2 \cap B_{mc}), \\ m_3 &= \text{MEAN}(B_3 \cap B_{mc}), m_4 = \text{MEAN}(B_4 \cap B_{mc}). \end{aligned}$$

$m_{mc}$  은 움직임 보상된 블록의 평균값을 의미하고,  $m_1, m_2, m_3, m_4$  는 각각 4개의 참조 블록에서의 국부영역에 대한 평균값을 의미한다.

또한 움직임 보상된 블록의 DC 값은 DCT영역에서 다음과 같은 수식으로 계산되어 질 수 있다.

$$m_{mc} = \frac{1}{64} \left[ \mathbf{T}_h \mathbf{X}_1 \mathbf{T}_w^T + \mathbf{T}_h \mathbf{X}_2 \mathbf{H}_w^T + \mathbf{H}_h \mathbf{X}_3 \mathbf{T}_w^T + \mathbf{H}_h \mathbf{X}_4 \mathbf{H}_w^T \right] \quad (2)$$

여기서

$$\mathbf{H}_x = \mathbf{A} \mathbf{U}_x \mathbf{C}^T, \mathbf{T}_x = \mathbf{A} \mathbf{L}_x \mathbf{C}^T, \mathbf{U}_x = \begin{bmatrix} \mathbf{I}_x & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \quad (3)$$

$$\mathbf{L}_x = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_x \end{bmatrix}, \mathbf{A} = [1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1]$$

$\mathbf{C}$  는  $8 \times 8$  DCT 행렬이고  $\mathbf{X}_i$  는  $B_i (1 \leq i \leq 4)$  의 DCT 계수들이다. 식 (2)에서 보면, 완전한 DC 영상을 얻기 위해서는 움직임 보상을 위해 완전한 해상도의 이전 프레임이 필요함을 알 수 있다.

## 2.2. DC 영상 추출을 위한 Zero-order 및 First-order 근사기법

Zero-order 근사기법에서는  $B_1, B_2, B_3, B_4$  중에서 움직임 보상 블록과 가장 넓은 중첩영역을 가지는 블록의 DC 값을 움직임 보상된 블록의 DC 값,  $m_{mc}$ , 로 간주한다 (그림 1). 즉,

$$m_{mc} = \begin{cases} \hat{m}_1, & \text{if } h \cdot w \geq h \cdot (8-w), (8-h) \cdot w, \text{ and } (8-h) \cdot (8-w) \\ \hat{m}_2, & \text{if } h \cdot (8-w) \geq h \cdot w, (8-h) \cdot w, \text{ and } (8-h) \cdot (8-w) \\ \hat{m}_3, & \text{if } (8-h) \cdot w \geq h \cdot w, h \cdot (8-w), \text{ and } (8-h) \cdot (8-w) \\ \hat{m}_4, & \text{if } (8-h) \cdot (8-w) \geq h \cdot w, h \cdot (8-w), \text{ and } (8-h) \cdot w \end{cases} \quad (4)$$

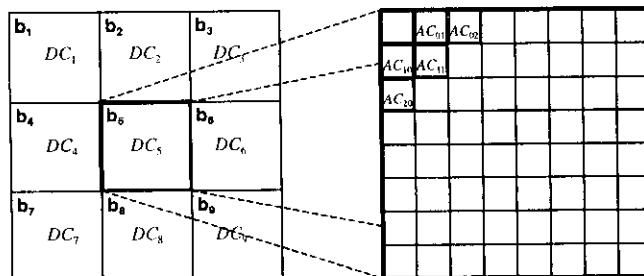
여기서

$$\begin{aligned} \hat{m}_1 &= \text{MEAN}(B_1), \hat{m}_2 = \text{MEAN}(B_2), \\ \hat{m}_3 &= \text{MEAN}(B_3), \hat{m}_4 = \text{MEAN}(B_4). \end{aligned} \quad (5)$$

반면 first-order 근사기법에서는 식 (1)에서의  $m_1, m_2, m_3, m_4$  가 각각  $\hat{m}_1, \hat{m}_2, \hat{m}_3, \hat{m}_4$  로 근사화 된다. 움직임 보상된 블록의 평균값은 다음과 같이 얻어진다.

$$m_{mc} = \frac{1}{64} \left[ \begin{matrix} h \cdot w \cdot \hat{m}_1 + h \cdot (8-w) \cdot \hat{m}_2 \\ + (8-h) \cdot w \cdot \hat{m}_3 + (8-h) \cdot (8-w) \cdot \hat{m}_4 \end{matrix} \right] \quad (6)$$

여기서  $\hat{m}_1, \hat{m}_2, \hat{m}_3, \hat{m}_4$  는 식 (5)와 같다.



8x8 DCT coefficients in the center block  $b_5$

그림 2. 중앙 블록  $b_5$  의 AC 계수 예측을 위해 사용되는 9 개의  $8 \times 8$  DCT 블록들

Zero-order 및 first-order 근사기법은  $m_{mc}$  를 추정하기 위해 참조 블록의 국부영역에 대한 DC 값을 구하는 대신 전체 블록의 DC 값을 사용한다. 이러한 접근은 알고리즘을 간단하게 하지만,  $m_{mc}$  의 부정확한 예측으로 인해 크기가 줄어든 DC 영상에서 좋지 않은 복원 화질을 얻게 된다.

### 3. 제안하는 DC 영상 추출 기법

DC 영상 추출을 주파수 영역에서 고려한다면, first-order 근사기법은 식 (2)에서 모든 AC 계수들을 0으로 간주한 것으로 볼 수 있다. 보다 정확한 DC 영상을 얻기 위해서는 정확한 AC 계수 정보들이 절대적으로 필요하지만 이것은 완전한 크기의 이전 프레임에 기반을 두기 때문에 추가적인 메모리와 연산능력을 요구한다. 제안하는 방법에서는 이러한 문제점을 극복하기 위해, 이차 평면 모델로부터 낮은 주파수에 해당하는 AC 계수들을 예측하는 방식을 사용한다.

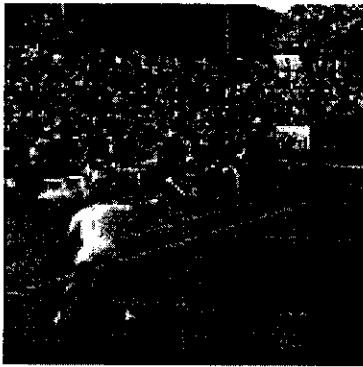
#### 3.1. 이차 평면 모델에 기반한 AC 예측기법

식 (2)에서  $X_i$  ( $1 \leq i \leq 4$ )의 5개의 낮은 주파수 AC 성분들을 예측하기 위해, 영상을 이차 평면 모델로 가정한다 [5]. 이때 이차 평면,

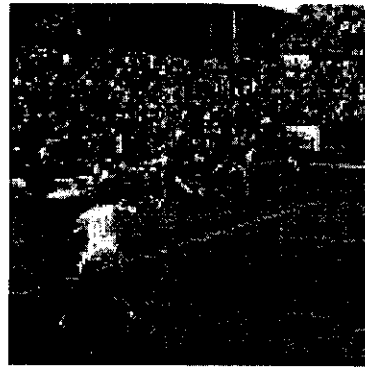
$$P(x,y) = a_{2,2}x^2y^2 + a_{2,1}x^2y + a_{1,2}xy^2 + a_{2,0}x^2 + a_{1,1}xy + a_{0,2}y^2 + a_{1,0}x + a_{0,1}y + a_{0,0} \quad (7)$$

은 그림 2에 묘사된 것처럼  $3 \times 3$  DC 값들에 대응된다.  $a_{2,2}, \dots, a_{0,0}$  는 이차 평면 모델로부터 구해지는 9개의 블록의 평균값들이,  $DC_1, \dots, DC_9$ 과 각각 일치한다는 조건하에서 얻어진다. 예측된 이차 평면 모델로부터 얻어지는 중앙 블록에 대해 DCT가 수행되면, 중앙 블록에 해당하는 AC 계수들은 주변의  $3 \times 3$  블록의 DC값들로부터 다음과 같이 구해진다.

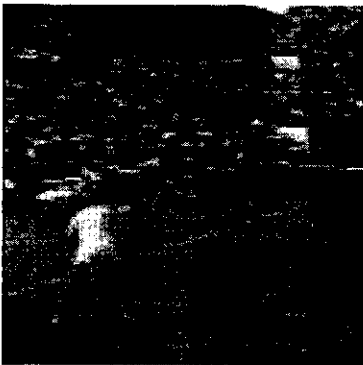
$$AC_{01} = (1.13885/8)(DC_4 - DC_8) \quad (8a)$$



(a)



(b)



(c)



(d)

그림 3. HDTV Tennis 영상의 23 번째 프레임에 대한 DC 영상 추출 결과; (a) 원영상, (b) zero-order 근사기법 [4], (c) first-order 근사기법[4], (d) 제안하는 방법.

$$AC_{10} = (1.13885/8)(DC_2 - DC_8) \quad (8b)$$

$$AC_{20} = (0.27881/8)(DC_2 + DC_8 - 2DC_5) \quad (8c)$$

$$AC_{11} = (0.16213/8)((DC_1 - DC_3) - (DC_7 - DC_9)) \quad (8d)$$

$$AC_{02} = (0.27881/8)(DC_4 + DC_6 - 2DC_5) \quad (8e)$$

AC 계수들은 인접한 DC 값들의 가중치가 가해진 선형 합으로 구해지므로 AC 계수들을 위한 추가적인 메모리는 필요 없다.

### 3.2. AC 예측을 이용한 DC 영상 추출

식 (2)에서 우리는 DCT영역에서 정확한 DC 영상을 추출하기 위한 방안을 유도했다. 이 방식이 정확한 AC 계수 값들을 요구하는 반면, 제안하는 방법에서는 몇 개의 AC 계수들만을 예측하는 방식을 사용하였다. 하나의 블록에 있어서 국부영역의 평균값은 낮은 주파수 성분에 의해 보다 큰 영향을 받으므로 zig-zag순으로 처음 5개의 계수들만 예측해서 사용한다. 제안된 DC 영상 추출방법은 다음과 같이 나타낼 수 있다.

$$\hat{m}_{dc} = \frac{1}{64} [T_h \hat{X}_1 T_w^T + T_h \hat{X}_2 H_w^T + H_h \hat{X}_3 T_w^T + H_h \hat{X}_4 H_w^T], \quad (9)$$

여기서  $\hat{X}_i$  는  $B_i (1 \leq i \leq 4)$ 의 추정된 DCT계수들이다.

제안된 방법에서는 DC 영상만 담을 수 있는 작은 크기의 메모리만 필요하다. 또한 AC 예측에도 불구하고 각 블록 당 곱하기 연산은 16번만 필요하다. 이는 식 (9)가 16개의 인접한 DC 값들의 가중치가 있는 선형 합으로 표현될 수 있고, 그 가중치 값들은 모든 경우에 대해서 사전에 미리 계산되어서 작은 메모리에 저장해 둘 수 있기 때문이다. 만약 식 (9)에서 모든 예측된 AC계수들을 0으로 간주된다면 first-order 근사기법과 동일하게 된다. 따라서 식 (9)는 DCT기반으로 압축된 동영상에서 DC 영상을 추출하는 보다 일반적인 표현 방법이라 볼 수 있다.

### 4. 모의 실험 결과

모의실험을 위해서 MPEG-1 동영상 압축기가 사용되었다. 1920x1080 크기의 두 개의 HDTV 영상, *Basketball*과 *Tennis*, 가 GOP 크기를 12로 해서 40Mbps로 부호화 되었다. 그림 3은 *Tennis* 영상에 대한 DC 영상 추출 결과를 보여준다. 예측 오차의 전과효과로 인해, zero-order 및 first-order 근사기법에서는 각각 눈에 거슬리는 부조화와 지나친 평탄화가 나타난다. 반면, 제안된 방법에서는 영상의 세밀함이 잘 보존되고 동시에 평탄한 영역에서 적절한 평탄화 효과가 나타나는 것을 볼 수 있다. 이것은 제안된 방식이 주요한 AC 계수들을 이용하고 있기 때문이다. 그림 4는 *Basketball* 영상에 대한 PSNR값들을 보여준다. 그림에서 알 수 있듯이,

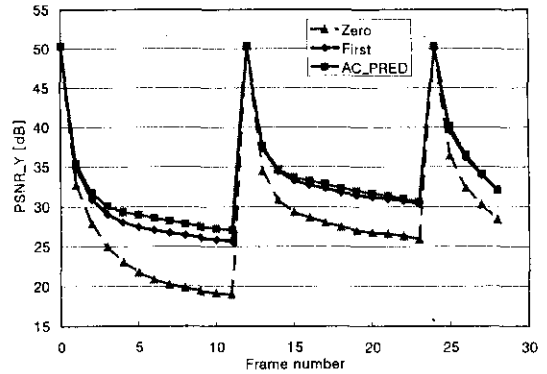


그림 4. HDTV *Basketball* 영상에 대한 PSNR 결과

제안된 방식이 first-order 근사기법에 비해 약간의 PSNR향상을 보여준다.

### 5. 결론

압축된 동영상의 빠른 해석과 효과적인 검색을 위한 새로운 DC 영상 추출 기법이 제안되었다. 제안된 방식에서는 이차 평면 모델을 기반으로 주요한 AC 성분을 예측함으로써 약간의 추가적인 연산량으로 보다 나은 화질을 얻을 수 있었다. 제안된 방식에서 AC 계수들은 이전 프레임의 복원된 DC 영상으로부터 얻어지기 때문에 zero-order 및 first-order 근사기법과 비교할 때 추가적인 메모리를 필요로 하지 않는다.

### 6. 참고문헌

- [1] N. Merhav and V. Bhaskaran, "Fast algorithms for DCT-domain image down-sampling and for inverse motion compensation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 7, pp. 468-476, June 1997.
- [2] B. L. Yeo and B. Liu, "Rapid scene analysis on compressed video," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 5, pp. 533-544, Dec. 1995.
- [3] Y. Nakajima, K. Ujihara, and A. Yoneyama, "Universal scene change detection on MPEG-coded data domain," in *Visual Communication and Image Processing*, SPIE, vol. 3024, pp. 992-1003, Feb. 1997.
- [4] B. L. Yeo and B. Liu, "On the extraction of DC sequence from MPEG compressed video," in *International Conference on Image Processing*, vol. II, pp. 260-263, Oct. 1995.
- [5] W. B. Pennebaker and J. L. Mitchell, *JPEG Still Image Data Compression*. New York: Van Nostrand Reinhard, 1993.