

## H.264의 Integer DCT 영역에서의 Intra-prediction 기법

\*안형진, 오형석, 김원하  
경희대학교 전자공학과

### A method for intra-prediction in the Integer DCT domain of H.264

\*Hyeong-Jin Ahn, Hyung-Suk Oh, Won-Ha Kim  
School of Electronics Engineering Kyunghee University

**Abstract** - 본 논문에서는 기존의 H.264/AVC의 spatial 영역에서 Intra prediction 기법과 달리 H.264/AVC에서 사용하는 Integer DCT 영역에서 Intra prediction 기법을 제안한다. 이를 위하여 Integer DCT 영역에서 Intra prediction을 수행하는 모든 과정을 matrix multiplication으로 표현하여 Intra prediction을 수행하는 matrix를 유도한다. Intra prediction을 수행하는 matrix를 각 모드에 알맞게 설계하고, 이 matrix를 Integer DCT 영역에서 사용할 수 있도록 orthogonal한 Integer matrix를 설계한다. 실험을 통하여 제안한 Integer DCT 영역에서 intra prediction 기법이 기존의 H.264/AVC의 spatial 영역에서 intra prediction 기법과 성능이 동일하면서 어떻게 matrix multiplication에 연산들을 포함시켜서 단순화 할 수 있는지를 보여 주었다. 또한 H.264/AVC에서 제공하는 intra prediction 각 모드에 대해 계산 복잡도를 분석하였다.

#### 1. 서 론

현재의 통신 환경은 유무선 이동, 방송망과 통신망의 융합, IP망을 이용한 IP convergence 등의 서비스를 가능케 하는 핵심 네트워크인 광대역 통합망이 나타나고 등 비균일 통신망이 서로 융합되고 있으며, DVD나 DTV, DMB 등의 시장이 커지고 있다. 이러한 추세에 맞추어 가자면, 서로 다른 코딩 환경에서 상황에 맞는 코덱을 사용하거나 보다 효율적이고 성능이 좋은 코덱을 사용해야 한다. 트랜스코딩은 서로 다른 시스템들과 디바이스들을 상호 이용이 가능토록 해주는 중요한 기술이다.

트랜스 코딩의 목적은 하나의 포맷으로부터 다른 포맷으로 멀티미디어 신호를 바꾸어 주는 것이다. 이와 같은 포맷 변환은 비트율, 프레임율, 코딩 표준 등에 의해 특징지어진다. 트랜스코딩이 transform 영역에서 이루어지면 완벽한 디코딩과 재 인코딩을 제공할 수 있고, 이는 계산적인 측면에서 상당한 이득이다. 이러한 요구에 부응하여 압축 효율을 최대화 시키는 H.264/AVC를 위한 트랜스 코딩 개발이 필요하다.

기존의 MPEG-2 트랜스코딩에는 없었던 H.264/AVC의 intra prediction 기법이 Integer DCT 영역에서 수행되어야 하며, 이를 위해서 근사화 DCT와 integer DCT의 matrix multiplication의 직교변환 성질과 PF값에 의한 양자화 등의 차이점을 파악하여 H.264/AVC의 intra prediction을 matrix multiplication으로 표현하는 방법에 대해 설명한다.

#### 2. 본 론

##### 2.1 H.264/AVC integer DCT

H.264/AVC에서 사용되는 정수 DCT는 DCT와 거의 동일한 transform coding gain을 내는 근사화 DCT에서 post scaling coefficient (PF) 행렬을 분리하여 구현된다.  $T_{N \times N}$ 과  $T_{N \times N}^{int}$ 을 각각 크기가 NxN인 근사화 DCT kernel 행렬과 정수 DCT kernel 행렬이라 한다. 근사화 DCT는 orthogonal 변환 행렬임으로  $T_{N \times N} \cdot T_{N \times N}^t = I_{N \times N}$ 이다.  $x_{N \times N}$ 를 크기가 NxN인 공간 영역의 데이터 블록이라 하고,  $X_{N \times N}$ 과  $X_{N \times N}^{int}$ 를 각각  $x_{N \times N}$ 을  $T_{N \times N}$ 과  $T_{N \times N}^{int}$ 로 변환한 DCT 계수 블록들이라 한다.  $x_{N \times N}$ 이 영상 데이터와 같이 정수 값의 블록이라면  $X_{N \times N}$ 는 실수 값의 DCT 블록이고  $X_{N \times N}^{int}$ 은 정수 값의 DCT 블록이 된다. 그러면 다음이 성립된다.

$$\begin{aligned} X_{N \times N} &= T_{N \times N} \cdot x_{N \times N} \cdot (T_{N \times N}^t)^t \\ &= \{T_{N \times N}^{int} \cdot x_{N \times N} \cdot (T_{N \times N}^{int})^t\} \otimes PF_{N \times N} \\ &= X_{N \times N}^{int} \otimes PF_{N \times N} \end{aligned} \quad (1)$$

여기서  $\otimes$ 은 element multiplication을 의미한다. 즉,  $X_{i,j}$ 와  $PF_{i,j}$ 를 각각  $X_{N \times N}^{int}$ 와  $PF_{N \times N}$ 의 (i,j)의 원소라 하면  $X_{N \times N}^{int} \otimes PF_{N \times N}$ 의 (i,j)의 원소는  $PF_{i,j} \cdot X_{i,j}$ 가 된다. H.264/AVC에서 사용하는 근사화 DCT  $T_{4 \times 4}^{AP}$ , 정수 DCT  $T_{4 \times 4}^{int}$ , PF 행렬  $PF_{4 \times 4}$ 는 다음과 같다.

$$T_{4 \times 4}^{AP} = \begin{bmatrix} a & a & a & a \\ b & c & -c & -b \\ a & -a & -a & a \\ c & -b & b & -c \end{bmatrix}, T_{4 \times 4}^{int} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & -2 \\ 1 & -1 & -1 & 1 \\ 1 & -2 & 2 & -1 \end{bmatrix} \quad (2)$$

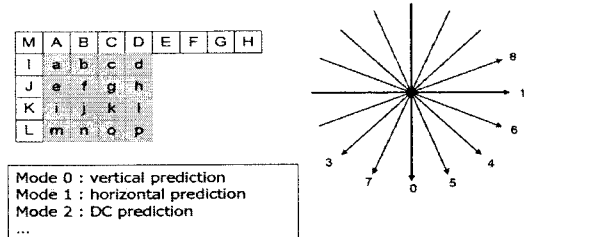
$$PF_{4 \times 4} = \begin{bmatrix} a^2 & \frac{ab}{2} & a^2 & \frac{ab}{2} \\ \frac{ab}{2} & \frac{b^2}{4} & \frac{ab}{2} & \frac{b^2}{4} \\ a^2 & \frac{ab}{2} & a^2 & \frac{ab}{2} \\ \frac{ab}{2} & \frac{b^2}{4} & \frac{ab}{2} & \frac{b^2}{4} \end{bmatrix} \quad (3)$$

여기서  $a = \frac{1}{2}, b = \sqrt{\frac{2}{5}}, c = \sqrt{\frac{1}{10}}$ 이다.

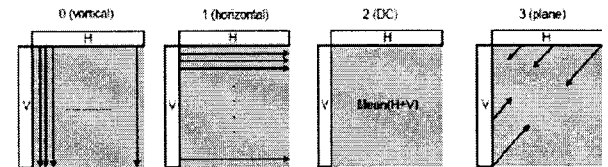
그러나, H.264/AVC에서 post-scaling factor  $PF_{N \times N}$ 는 정수 연산을 위하여 순방향 양자화에 통합된다.

##### 2.2 H.264/AVC Intra-prediction

H.264/AVC에서 Intra prediction 기법은 이웃한 화소들 간의 공간적인 상관관계를 이용한 화면 내 예측 기법이다. H.264/AVC에서는 휘도 샘플의 경우 블록 모드가 Intra4x4 모드일 경우 9개의 예측 모드를 정의하고 있고, Intra16x16 모드와 색차 성분에는 경우 4개의 예측 모드를 정의하고 있다.



<그림 1> 4x4 Intra-prediction



<그림 1> 16x16 및 색차성분에 대한 Intra-prediction

##### 2.3 H.264/AVC의 정수 DCT 영역에서 Intra-prediction

H.264/AVC에서 제공하는 각 intra prediction을 matrix multiplication과 직교 변환의 성질을 이용하여 기존에 spatial 영역에서 이루어지던 intra prediction을 Integer DCT 영역에서 할 수 있는 알고리즘을 각 모드 별로 다음과 같이 정의한다.

###### 2.3.1 Vertical mode

Vertical mode는 현재 부호화하고자 하는 블록의 이웃하는 이미 코딩된 위 블록의 4픽셀들(A to D)을 이용하여 예측하는 모드이다. 여기서 A 픽셀은 블록의 첫 column 부분의 4개의 픽셀에 채워지고 B 픽셀은 블록의 두 번째 column 부분의 4개의 픽셀에 채워지며, C, D 픽셀은 두 각각 블록 column에 해당하는 4개의 픽셀에 채워진다. 이와 같이 Vertical mode로 예측 블록이 생성되는 것을 다음과 같이 행렬식으로 표현할 수 있다.

$$\begin{bmatrix} A & B & C & D \\ A & B & C & D \\ A & B & C & D \\ A & B & C & D \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} X & X & X & X \\ X & X & X & X \\ X & X & X & X \\ A & B & C & D \end{bmatrix} \quad (4)$$

여기서 X는 "don't care" 픽셀이다.

하지만 H.264/AVC에서 사용하는 integer DCT는 직교변환 성질이 없기 때문에 integer DCT 영역에서 Intra-prediction을 하기 위해서는 다음과 같이 integer DCT를 인위적으로 변형시켜 직교변환 성질을 가지도록 해야 한다.

$$T_{forw}^t \cdot A \cdot T_{orw} = I$$

$$A = (T_{inv}^t \cdot T_{orw}^t)^{-1} \cdot T_{inv}^t \cdot (T_{orw}^t \cdot T_{inv}^t)^{-1} \quad (5)$$

A는 integer DCT의 직교변환 성질을 만들어 주기 위한 matrix이므로 A를 이용하여 Vertical 예측 모드 행렬식을 표현하면 다음과 같다.

$$T_{4 \times 4}^{int} \begin{Bmatrix} A & B & C & D \\ A & B & C & D \\ A & B & C & D \\ A & B & C & D \end{Bmatrix} = T_{4 \times 4}^{int} \begin{Bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{Bmatrix} \cdot \begin{Bmatrix} X & X & X & X \\ X & X & X & X \\ X & X & X & X \\ A & B & C & D \end{Bmatrix} \quad (6)$$

$$= T_{4 \times 4}^{int} \begin{Bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{Bmatrix} \cdot A \cdot T_{4 \times 4}^{int} \begin{Bmatrix} X & X & X & X \\ X & X & X & X \\ X & X & X & X \\ A & B & C & D \end{Bmatrix}$$

$$= \begin{bmatrix} 1 & -0.8 & 1 & -0.6 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \cdot T_{4 \times 4}^{int} \begin{Bmatrix} X & X & X & X \\ X & X & X & X \\ X & X & X & X \\ A & B & C & D \end{Bmatrix}$$

### 2.3.2 Horizontal mode

Horizontal mode는 현재 부호화하고자 하는 블록의 이웃하는 이미 코딩된 왼쪽 블록의 4픽셀(L to L)을 이용하여 예측하는 모드이다. 여기서 1픽셀은 블록의 첫 row 부분의 4개의 픽셀에 채워지고 J 픽셀은 블록의 두 번째 row 부분의 4개의 픽셀에 채워지며, K, L 픽셀도 각각 블록 row에 해당하는 4개의 픽셀에 채워진다. 이와 같이 Horizontal mode로 예측 블록이 생성되는 것을 위와 같은 방법으로 integer DCT 영역에서 표현하면 다음과 같다.

$$\begin{Bmatrix} I & I & I & I \\ J & J & J & J \\ K & K & K & K \\ L & L & L & L \end{Bmatrix} = \begin{Bmatrix} X & X & X & I \\ X & X & X & J \\ X & X & X & K \\ X & X & X & L \end{Bmatrix} \cdot \begin{Bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{Bmatrix} \quad (7)$$

$$T_{4 \times 4}^{int} \begin{Bmatrix} I & I & I & I \\ J & J & J & J \\ K & K & K & K \\ L & L & L & L \end{Bmatrix} = T_{4 \times 4}^{int} \begin{Bmatrix} X & X & X & I \\ X & X & X & J \\ X & X & X & K \\ X & X & X & L \end{Bmatrix} \cdot \begin{Bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{Bmatrix} \quad (8)$$

$$= T_{4 \times 4}^{int} \begin{Bmatrix} X & X & X & I \\ X & X & X & J \\ X & X & X & K \\ X & X & X & L \end{Bmatrix} \cdot A \cdot T_{4 \times 4}^{int} \begin{Bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{Bmatrix}$$

$$= T_{4 \times 4}^{int} \begin{Bmatrix} X & X & X & I \\ X & X & X & J \\ X & X & X & K \\ X & X & X & L \end{Bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 & 0 \\ -0.8 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ -0.6 & 0 & 0 & 0 \end{bmatrix}$$

### 2.3.3 DC mode

DC mode는 현재 부호화하고자 하는 블록의 이웃하는 위쪽 4픽셀과 왼쪽 4픽셀의 평균값으로 예측 블록을 생성하는 모드이다. 위와 같은 방식으로 행렬식을 표현하면 다음과 같다.

$$\left\{ \frac{A+B+C+D+I+J+K+L}{8} \right\} \cdot \begin{Bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{Bmatrix} \quad (9)$$

$$T_{4 \times 4}^{int} \cdot \left\{ \frac{A+B+C+D+I+J+K+L}{8} \right\} \cdot \begin{Bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{Bmatrix} \quad (10)$$

$$= \frac{1}{8} \cdot T_{4 \times 4}^{int} \begin{Bmatrix} A & B & C & D \\ A & B & C & D \\ A & B & C & D \\ A & B & C & D \end{Bmatrix} \cdot \begin{Bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{Bmatrix} + \frac{1}{8} \cdot T_{4 \times 4}^{int} \begin{Bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{Bmatrix} \cdot \begin{Bmatrix} I & I & I & I \\ J & J & J & J \\ K & K & K & K \\ L & L & L & L \end{Bmatrix}$$

$$= \frac{1}{8} \cdot T_{4 \times 4}^{int} \begin{Bmatrix} A & B & C & D \\ A & B & C & D \\ A & B & C & D \\ A & B & C & D \end{Bmatrix} \cdot A \cdot T_{4 \times 4}^{int} \begin{Bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{Bmatrix}$$

$$+ \frac{1}{8} \cdot T_{4 \times 4}^{int} \begin{Bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{Bmatrix} \cdot A \cdot T_{4 \times 4}^{int} \begin{Bmatrix} I & I & I & I \\ J & J & J & J \\ K & K & K & K \\ L & L & L & L \end{Bmatrix}$$

$$= \frac{1}{8} \cdot T_{4 \times 4}^{int} \begin{Bmatrix} A & B & C & D \\ A & B & C & D \\ A & B & C & D \\ A & B & C & D \end{Bmatrix} \cdot \begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + \frac{1}{8} \cdot \begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \cdot T_{4 \times 4}^{int} \begin{Bmatrix} I & I & I & I \\ J & J & J & J \\ K & K & K & K \\ L & L & L & L \end{Bmatrix}$$

이와 비슷한 방법으로 남은 intra4x4 모드들과 intra16x16 모드 및 색차 신호 intra8x8 모드들에 대해서도 쉽게 유도 할 수 있다.

### 2.4 실험

본 논문에서는 Integer DCT 영역에서 intra prediction한 영상과 spatial 영역에서 intra prediction한 영상을 이미지의 객관적인 화질 측정 방법인 PSNR을 이용하여 측정할 경우 차이가 미세하므로 주관적인 화질을 이용하여 측정하였고 두 이미지의 차이 값을 보여준다. 이때 QP값은 Qstep size가 1이 되는 4로 놓고 실험하였다. 그리고 프리딕션 샘플을 얻기 위해서 요구되는 연산들의 수를 제한된 알고리즘의 4x4 intra prediction의 4가지 모드에 대해서 계산상의 복잡도를 보여준다.



(a)



(b)



(c)

<그림 1> (a) H.264/AVC의 spatial 영역에서 intra prediction한 이미지. (b) 본 논문에서 제안하는 기법을 적용하여 intra prediction한 이미지. (c) (a)와(b)와의 차이 이미지.

<표 1> 4x4 intra prediction의 계산상의 복잡도

Moed	Multiplication	Addition
Vertial	8	12
Horizontal	8	12
DC	1	1
Diagonal_Down_Left	168	136

### 3. 결 론

이 논문은 H.264/AVC의 기존의 spatial 영역에서 intra prediction 기법을 변환하여 Integer DCT 영역에서 Intra prediction을 하는 방법에 관한 접근을 논하였다. 논문의 접근 방법은 근사화 DCT와는 다른 Integer DCT의 matrix multiplication의 기본 성질을 이용하여 직교변환 성질을 가지게 한 다음 H.264/AVC의 integer DCT 영역에 적용 시킨 것이다. Integer DCT 영역에서 intra prediction을 matrix multiplication을 이용해서 함으로써 단순화 시킬 수 있고, 완벽한 디코딩과 재인코딩 과정을 생략할 수 있으며, 예를 들어 MPEG2 to H.264/AVC 포맷 변환과 같은 transform 영역에서 트랜스코딩을 수행하는데 실용적으로 많은 도움이 될 것이다.

### 3. 감사의 글

본 연구는 과학재단 특정기초 연구사업(R01-2005-000-11054-0)의 지원으로 수행되었음.

### [참고 문헌]

- [1] P. A. A. Assuncao and M. Ghanbari, "Transcoding of MPEG-2 video in the frequency domain," in *IEEE Int. Conf. Acoustics, Speech, Signal Processing, ICASSP'97*, Munich, Germany, Apr. 1997, vol. 4.
- [2] T. Shanableh, and M. Ghanbari, "Hybrid DCT/pixel domain architecture for heterogeneous video transcoding," *Signal Processing: Image Commun.*, vol. 18, no. 8, pp. 601-620, Sep. 2003.
- [3] Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification (ITU-T Rec. H.264 | ISO/IEC 14496-10/AVC), 2003.
- [4] C. Chen, P.-H. Wu, and H. Chen, "MPEG-2 to H.264 transcoding," *Picture Coding Symposium*, Dec. 2004.
- [5] C. Chen, P.-H. Wu, and H. Chen, "Transform-Domain Intra Prediction for H.264," *IEEE International Symposium on Circuits and Systems (ISCAS)*, May, 2005.