H.264/AVC 압축 도메인에서의 프레임 resizing 방법

Frame resizing scheme in H.264/AVC compressed domain

*오 형 석, **김 원 하
(Hyung Suk Oh, Won Ha Kim)

Abstract - Image resizing is to change an image size by upsampling or downsampling of a digital image. Most still images and video frames are given in a compressed domain on digital media. Image resizing of a compressed image can be performed in a spatial domain via decompression or recompression. In general, resizing of a compressed image in a compressed domain is much faster than that in a spatial domain. In this paper, we propose an approach to resize images in the integer discrete cosine transform (DCT) domain, which exploits the multiplication-convolution property of DCT.

Key Words : H.264/AVC, resizing, compressed domain

1. 개요

현재의 통신환경은 유무선 연동, 방송망과 통신망의 융합, IP망을 이용한 IP Convergence등의 서비스를 가능케 하는 핵심 네트워크인 광대역 통합망(BcN :Broadband Convergence Network)이 나타나는 등 비균일 통신망이 서로 융합되고 있으며, 앞으로도 이러한 추세는 더욱 가속화될 것이다. 이와 같은 비균일 통신 환경에서 디지털 융합(Digital Convergence)의 추세에 맞추어 가자면 동영상 전송도 범용 미디어 통신 (Universal Media Access, UMA)이 되도록 해야 한다. 따라서 동영상의 부호화 방법은 압축효율을 최대화 시키는 것뿐 만 아니라 다양한 단말기 및 변화하는 통신 환경에 친화적으로 대응할 수 있는 image resizing 기법이 도입되어야 한다. 이러한 요구에 부응하여 압축효율을 최대화 시키는 H.264/AVC compressed domain에서 resizing기법은 방송 및 감시 시스템 등 많은 멀티미디어 관련 분야에서 활용 가치가 높다.

본 논문에서 H.264/AVC integer DCT의 설명을 2장에서 설명하며, 3장에서는 DCT의 multiplication convolution특성을 설명한다. compressed domain에서 다운샘플링, 업샘플링을 하는 제안 기법은 4, 5장에서 설명한다.

2. H.264/AVC integer DCT

일반적으로 기존의 DCT는 JPEG과 MPEG에서 블록 변환 기법으로 사용되었다. DCT는 길이 N의 샘플 x개를 주파수 도메인에서 새로운 변환 계수를 갖는 X로 만드는 것이다. kth 열과 nth 행을 갖는 변환식 T를 다음과 같이 정의한다.

$$T_{kn} = T(k,n) = c_k\sqrt{\frac{2}{N}}\, cos\left(\frac{(2n+1)k\pi}{2N}\right) \quad (1)$$

$$c_0 = \sqrt{2},\ k=0$$
$$c_k = 1,\quad k>0$$

DCT의 단점은 $T(k,n)$의 값들이 실수란 점이다. 따라서 DCT의 단점을 보완한 H.264/AVC에서 새로이 도입된 정수 DCT는 직교특성을 유지하면서 $T(k,n)$의 값들을 가장 근접한 정수로 바꾸는 것이다.[1][3][4] 따라서 기존의 DCT와 정수 DCT를 다음과 같이 표현할 수 있다.

$$T=\begin{bmatrix} a & a & a & a \\ b & c & -c & -b \\ a & -a & -a & a \\ c & -b & b & -c \end{bmatrix}$$

$$a=\frac{1}{2},\ b=\sqrt{\frac{1}{2}}\,cos\left(\frac{\pi}{8}\right),\ c=\sqrt{\frac{1}{2}}\,cos\left(\frac{3\pi}{8}\right) \quad (2)$$

이 행렬은 다음과 같은 정수 형태로 변환될 수 있다.

$$H=\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & d & -d & -1 \\ 1 & -1 & -1 & 1 \\ d & -1 & 1 & -d \end{bmatrix}$$

$$d=\frac{1}{2} \quad (3)$$

이 변환은 4x4 DCT를 근사화 시킨 것이지만, 계수 d와 b의 변화로 인해 새로운 변환의 결과는 4x4 DCT의 결과와 동일하지 않다. 하지만 H.264/AVC 코덱에서 정수 DCT는 DCT와 거의 동일한 압축 성능을 나타내며, 덧셈, 뺄셈 그리고 쉬프트 연산만을 사용하여 정수 연산을 수행할 수 있다는 장점을 가지고 있다.

저자 소개
* 오형석 : 경희대학교 전자공학과 석사과정
** 김원하 : 경희대학교 전자정보학부 부교수

3. DCT의 Multiplication-Convolution 특성

Transform domain에서 image resizing의 효율적인 방법을 표현하기 위해서 spatial 과 transform domain의 관계를 알아야 한다. 길이 N을 가진 x(n)의 DCT를 다음과 같은 식으로 표현한다.

$$X(k) = C_{2e}\{x(n)\} = 2\sum_{n=0}^{N-1} x(n)\cos\frac{k\pi(n+\frac{1}{2})}{N} \quad \text{for } 0 \leq k, n < N \tag{4}$$

$$x(n) = C_{2e}^{-1}\{X(k)\} = \frac{1}{N}\sum_{k=0}^{N-1} u(k)X(k)\cos\frac{k\pi(n+\frac{1}{2})}{N} \quad \text{for } 0 \leq k, n < N \tag{5}$$

$u(k)$는 다음과 같이 정의 한다.

$$u(k) = \begin{cases} \frac{1}{2}, & k=0 \\ 1, & 1 \leq k < N \end{cases} \tag{6}$$

DCT는 다음과 같은 multiplication-convolution 특성을 갖는다.

$$x(n) \times w(n) = C^{-1}{}_{2e}\{C_{2e}\{x(n)\} \otimes_s C_{2e}\{w(n)\}\} \tag{7}$$

$\otimes_s$는 symmetric convolution을 나타내는 기호이며, x(n)과 w(n)는 spatial domain에서의 신호이다. multiplication-convolution 특성은 잘 알려진 Discrete Fourier Transform(DFT)의 convolution 특성이다. DCT는 symmetric 확장된 신호의 DFT로 유도된다.

길이 N인 신호 x(n), symmetric 확장된 신호를 $\bar{x}(n)$이라 하면 $\bar{x}(n)$을 다음과 같이 정의 할 수 있다.[2]

$$\bar{x}(n) = \begin{cases} x(n) & 0 \leq n < N \\ x(2N-n-1) & N \leq n < 2N \end{cases} \tag{8}$$

n=1/2인 대칭점에서 신호 $\bar{x}(n)$의 DFT는 다음과 같이 나타낸다.

$$\overline{X_{DFT}}(k) = \sum_{n=0}^{2N-1} \bar{x}(n)\exp\left(\frac{-j2\pi k\left(n+\frac{1}{2}\right)}{2N}\right) = \sum_{n=0}^{N-1} 2\bar{x}(n)\cos\left(\frac{\pi k}{N}\left(n+\frac{1}{2}\right)\right) \quad 0 \leq k < 2N \tag{9}$$

위의 식에서 DCT X(k)는 다음과 같이 나타낸다.

$$X(k) = C_{2e}\{x(n)\} = \overline{X_{DFT}}(k) = 2\sum_{n=0}^{N} x(n)\cos\left(\frac{\pi k}{N}\left(n+\frac{1}{2}\right)\right) \quad 0 \leq n, k < N \tag{10}$$

위의 두 식에서 다음과 같은 특성을 알 수 있다.

$$\overline{X_{DFT}}(k) = -\overline{X_{DFT}}(2N-k),\ 0 \leq k < 2N$$

$$X_e(k) \begin{cases} X(k), & 0 \leq k < N \\ 0, & k = N \\ -X(2N-k), & N < k < 2N \end{cases} \tag{11}$$

따라서 symmetric convolution 특성을 다음과 같이 나타낼 수 있다.

$$X(k) \otimes_s W(k) = \sum_{i=0}^{k} X_e(i)W_e(k-i) - \sum_{i=k+1}^{2N-1} X_e(i)W_e(k-i+2N), 0 \leq k < N \tag{12}$$

본 논문의 image resizing 기법은 식 (7)과 (12)를 이용한 multilplication-convolution 특성을 이용하였다.

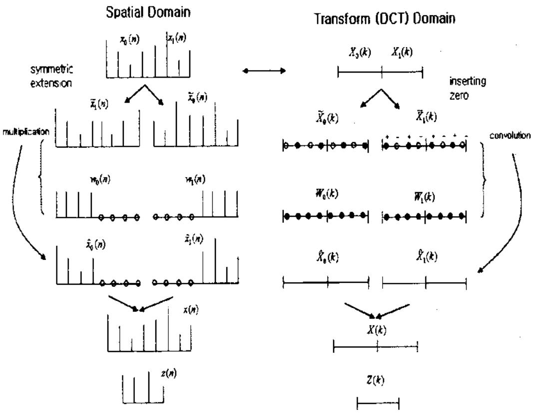
4. Downsampling in DCT domain

본 논문에서 다운샘플링과 업샘플링의 설명은 1차원으로 설명한다.

<그림1>은 다운샘플링 과정을 spatial 과 DCT domain에서 나타낸 그림이다.

Spatial domain에서 $w(n)$과 $x(n)$의 multiplication 연산은 DCT domain에서 식(7)번과 같이 symmetric convolution과 같다. 결과적으로 DCT domain에서 $X(k)$는 $\widetilde{X_0}(k)$와 $\widetilde{X_1}(k)$의 다음과 같은 관계식으로 얻을 수 있다.

$$X(k) = \widetilde{X_0}(k) \otimes_s W_0(k) + \widetilde{X_1}(k)W_1(k) \tag{13}$$



<그림1> Transform domain에서 Downsampling 과정

위의 관계식 (13)의 $\widetilde{X_0}(k) \otimes_s W_0(k)$를 식(14)로 표현한다.

$$\widetilde{X_0}(k) \otimes_s W_0(k), \text{for} 0 \leq k < 8 \Leftrightarrow W_0^{d2}X_0 \tag{14}$$

W_0^{d2}, W_1^{d2}는 symmetric convolution filter이다.

또한 식(13)의 두 번째 관계식을 식(14)와 같이 표현하면,

$$\widetilde{X_1}(k) \otimes_s W_1(k), \text{for} 0 \leq k < 8 \Leftrightarrow W_1^{d2}X_1 = SW_0^{d2}SX_1$$

$$(15)$$

$$S = diag(1,-1,1,-1,1,-1,1,-1) \tag{16}$$

따라서 다운샘플링된 z(k)는 다음과 같이 표현할 수 있다.

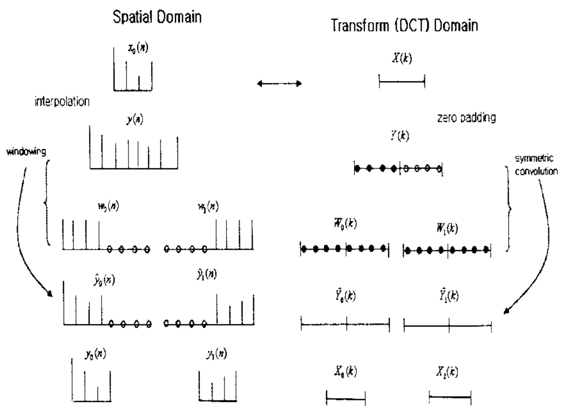
$$Z(k) = GW_0^{d2}FX_0 + GSW_0^{d2}SFX_1 \tag{17}$$

$$F = diag(4\sqrt{2},4,4,4,4,4,4,4) \tag{18}$$

$$G = diag(\frac{1}{4\sqrt{2}},\frac{1}{4},\frac{1}{4},\frac{1}{4},\frac{1}{4},\frac{1}{4},\frac{1}{4},\frac{1}{4}) \tag{19}$$

5. Upsampling in DCT domain

일반적으로 DCT domain에서의 업샘플링은 <그림2>와 같이 고주파영역에 0을 넣음으로써 해결할 수 있다.



<그림2> Transform domain에서 Upsampling 과정

multiplication-convolution 관계를 이용하여 다운샘플링과 마찬가지로 관계식을 풀이할 수 있다.

$$X_0(k) = GW_0^{d2}FX$$

$$X_1(k) = GSW_1^{d2}FX \tag{20}$$

6. 결론

이 논문에서는 transform domain에서 images를 resizing 하는 방법에 관한 접근을 논하였다. 논문의 접근 방법은 DCT의 multiplication-convolution의 기본 성질을 이용하여 H.264/AVC의 integer DCT에 적용시킨 것이다. H.264/AVC의 integer DCT를 이용하여 다운 샘플링 혹은 업샘플링을 함으로써 서로 다른 크기의 이미지들의 화질을 보장할 수 있을 것이다. 또한 시간적 효율, 즉 계산량에서도 더 빠를 것이며, spatial domain에서 resizing하는 것보다 더 낮은 복잡도를 가지게 되는 장점이 있다.

감사의 글

본 연구는 경기도 D-TV 연구사업(다 매체를 지원하는 양방향 D-TV 시스템 개발)의 지원으로 수행되었음.

참 고 문 헌

[1] H. S. Malvar, A. Hallapuro, M. Karczewicz and L. Kerofsky, "Low-Complexity Transform and Quantization in H.264/AVC" IEEE Trans. Circuits. Syst., Video Technol., vol. 13, No. 7 July 2003.

[2] A. K. Jain, Fundamentals of Digital Image Processing. Englewood Cliffs, NJ: Prentice-Hall, 1989.

[3] Iain E.G. Richardson, " H.264 and MPEG-4 Video Compression", Wiley, 2004.

[4] 정제창역, "H.264/AVC 비디오 압축 표준", 홍릉과학출판사, 2005