

변경된 IDCT를 이용한 영상의 축소 및 확대 기법

서성주, 이명희, 설상훈

고려대학교 전자공학과

sjsuh@mpeg.korea.ac.kr, mhlee@mpeg.korea.ac.kr, sull@mpeg.korea.ac.kr

Image Resizing in the Compressed Domain Using Modified IDCT

Sungjoo Suh, Myung-Hee Lee and Sanghoon Sull

Dept. of Electronics Engineering, Korea University

요 약

영상압축기술의 발전으로 인하여 디지털 영상은 대부분 압축된 형태로 사용된다. 이러한 압축된 영상을 축소 및 확대하는 것은 네트워크 환경에서의 대역폭에 따른 축소된 영상 전송과 디스플레이 장치에 맞는 크기로의 영상 조절 등 다양한 응용에 사용 가능하다. 가장 대표적인 압축 방식은 DCT를 이용해서 영상을 부호화하는 것이다. 최근 DCT로 부호화된 영상에 대해 DCT 영역에서 직접 축소한 후 다시 DCT 영역에서 확대해 원래 크기의 DCT로 부호화된 영상을 얻는 방법에 대한 연구가 진행되어왔다. 이러한 연구의 주 관심사는 최종적으로 부호화된 결과 영상의 화질을 개선하는 것이다. 본 논문에서는 DCT로 부호화된 영상을 축소하기 위해, 변경된 IDCT 방법을 제안한다. 그리고 변경된 IDCT 방법을 사용해서 DCT 영역에서 직접 축소된 영상을 얻는 방법과 이에 대응하는 확대된 영상을 얻는 새로운 방법을 제시한다. 제시된 영상 축소 방법과 확대 방법을 같이 사용함으로써 DCT 영역에서 축소 후 확대된 영상은 가장 최근에 제안된 방법들보다 높은 PSNR 값을 나타낸다.

I. 서 론

JPEG, MPEG 등의 많은 영상압축 표준들은 DCT를 이용해서 영상을 부호화한다. DCT 영역에서 직접 영상을 처리하는 것은 복호와 부호에 따른 변환 과정에서 발생하는 많은 양의 데이터를 줄일 수 있는 장점이 있다. DCT로 부호화된 영상의 크기를 조절하는 일반적인 방법은 먼저 부호화된 영상을 복호하고, 공간영역에서 영상의 크기를 조절한 후, 크기가 조절된 영상을 다시 부호화한다. 그러나 이러한 공간영역에서의 접근은 많은 계산량을 요구한다.

최근 DCT 영역에서 영상의 크기를 빠르게 조절하는 방법이 Dugad [1]에 의해 제안되었다. Dugad의 방법에서 원 영상의 DCT 블록들은 축소된 영상의 DCT 블록들로 DCT 영역에서 직접 변환된다. 이와 유사하게 원 영상(또는 축소된 영상)의 DCT 블록들은 확대된 영상의 DCT 블록들로 DCT 영역에서 직접 변환된다. 더욱이, [1]에서 제안된 방법으로 영상을 축소 후 확대하면 원 영상의 모든 저주파성분을 보존하게 된다. 그러나 확대된 영상에서 모든 고주파성분들이 없어지게 되는 단점이 있다. Mukherjee [2]는 서브밴드 DCT를 이용한 DCT 영역에서의 영상 크기 조절 방

법을 제안하였다. Mukherjee는 DCT 영역에서의 영상 확대를 위해 16×16 DCT를 사용하였다. 그 결과 Dugad의 방법과 달리 확대된 영상의 고주파 성분이 존재하게 된다. 그러나 [2]에서 제안된 방법으로 영상을 축소 후 확대하게 되면 원 영상의 저주파 성분이 그대로 보존되지 못하는 문제점이 있다.

본 논문에서는 [2]에서 제안된 방법을 기초로 하여 변경된 IDCT를 사용하여 DCT 영역에서의 영상 축소를 위한 방법과 영상 확대 방법을 제안한다.

본 논문의 구성은 다음과 같다. 2장에서는 DCT 영역에서의 영상 축소를 위해 사용되는 변경된 IDCT 방법이 설명되고, 이를 이용한 영상의 축소 및 확대 기법이 3장에서 기술된다. 아울러 4장에서는 본 논문에서 제안된 방법과 기존의 방법과의 비교 실험결과가 제시되며, 마지막으로 5장에서는 결론을 맺는다.

II. 다운샘플링에서의 변경된 IDCT

공간 영역에서 영상을 축소하는 방법은 일반적으로 영상을 저역필터를 통과시킨 후에 다운샘플링을 통해서 이루어진다. 영상크기를 축소하기 위해 일반적으로는 선형보간

법(주변 픽셀들의 평균)을 사용한다. 선형보간법에서 일차원인 경우를 생각해 보면, 우선 $n=0,1,\dots,N-1$ 인 $N(N$ 은 짝수)개의 픽셀들을 $x(n)$ 이라 하고, $n=0,1,\dots,N/2-1$ 인 $N/2$ 개의 줄어든 픽셀들을 $x_R(n)$ 이라고 하자. 그러면 $x_R(n)$ 은 다음과 같이 표현될 수 있다.

$$x_R(n) = \frac{1}{2}[x(2n) + x(2n+1)], n=0, 1, \dots, \frac{N}{2}-1 \quad (1)$$

$X(k)(k=0,1,\dots,N-1)$ 을 $x(n)$ 의 DCT값이라고 놓으면, $X(k)$ 의 IDCT는 다음과 같이 정의된다.

$$x(n) = \sum_{k=0}^{N-1} \alpha(k) X(k) \cos\left(\frac{(2n+1)k\pi}{2N}\right) \quad (2)$$

$$\alpha(k) = \begin{cases} \sqrt{\frac{1}{N}}, & k=0 \\ \sqrt{\frac{2}{N}}, & \text{otherwise} \end{cases}$$

식(2)를 DCT항을 포함한 $x(2n)$ 과 $x(2n+1)$ 으로 표현하고, 간단한 수학 공식을 적용하여 다음과 같이 나타낼 수 있다.

$$\begin{aligned} x_R(n) &= \frac{1}{2} \left[\sum_{k=0}^{N/2-1} \alpha(k) X(k) \cos\left(\frac{(4n+1)k\pi}{2N}\right) \right. \\ &\quad \left. + \sum_{k=0}^{N/2-1} \alpha(k) X(k) \cos\left(\frac{(4n+3)k\pi}{2N}\right) \right] \\ &= \sum_{k=0}^{N/2-1} \alpha(k) X(k) \cos\left(\frac{k\pi}{2N}\right) \cos\left(\frac{(2n+1)k\pi}{N}\right) \quad (3) \end{aligned}$$

일반적으로 대부분의 에너지는 저주파영역에 밀집해 있기 때문에 식(3)에서 고주파 성분은 무시할 수 있다. 그러면, $x_R(n)$ 은

$$x_R(n) \approx \sum_{k=0}^{N/2-1} \alpha(k) Y(k) \cos\left(\frac{(2n+1)k\pi}{N}\right) \quad (4)$$

$$Y(k) = X(k) \cos\left(\frac{k\pi}{2N}\right), k=0, 1, \dots, N/2-1$$

와 같이 근사될 수 있다.

그러므로 일차원인 경우에 영상의 크기 축소는 우선, DCT계수 $X(k)$ 를 변형시켜 저주파수 영역에서 변경된 새로운 DCT계수 $Y(k)$ 를 만든 후, 새로 만들어진 $Y(k)$ 에 $N/2$ IDCT를 수행한다. 이차원으로서의 확장은 일차원인 경우를 바탕으로 다음과 같이 쉽게 할 수 있다. $x(m, n)$ 를 $N \times N$ 블록의 픽셀이라고 하고, $x_R(m, n)$ 을 가로, 세로 $1/2$ 만큼씩 크기가 줄어든 블록의 픽셀이라고 하자. 이차원의 경우에 식(4)는

$$\begin{aligned} x_R(m, n) &= \frac{1}{4} [x(2m, 2n) + x(2m, 2n+1) \\ &\quad + x(2m+1, 2n) + x(2m+1, 2n+1)] \\ &\approx \sum_{u=0}^{N/2-1} \sum_{v=0}^{N/2-1} \alpha(u, v) Y(u, v) \cos\left(\frac{(2m+1)u\pi}{N}\right) \\ &\quad \times \cos\left(\frac{(2n+1)v\pi}{N}\right), m, n=0, 1, \dots, N/2-1 \quad (5) \\ \alpha(u, v) &= \alpha(u)\alpha(v), \\ Y(u, v) &= X(u, v) \cos\left(\frac{u\pi}{2N}\right) \cos\left(\frac{v\pi}{2N}\right) \end{aligned}$$

로 확장된다.

DCT로 부호화된 영상의 크기를 줄이려면 $N \times N$ 블록에

서 식(5)와 같은 방법으로 각각의 DCT 계수를 변형한 후에 $N/2 \times N/2$ IDCT를 적용해서 구할 수 있다. 크기를 줄이는데 사용되는 변경된 IDCT는 $N \times N$ DCT 블록 위쪽 $N/2 \times N/2$ 부분을 제외한 모든 DCT계수를 0으로 대치하고 $N \times N$ IDCT를 적용한 후에 선형보간을 한것으로 해석될 수 있고, [3]에서의 서브밴드 IDCT의 근사화라고 해석될 수도 있다. 따라서, 변경된 IDCT를 적용하여 크기를 줄인 영상은 저역필터를 통과하고 다운 샘플링된 영상이라 할 수 있다.

III. 영상크기 조정 방법

이번 장에서는 본 논문에서 제안된 변경된 IDCT 방법으로 직접 DCT영역에서 영상을 축소하고 확대하는 방법을 설명한다. 설명의 편리성을 위해 b 를 공간영역에서의 8×8 블록, B 를 DCT 영역에서의 8×8 블록이라 한다.

1. 영상의 축소 방법

DCT영역의 영상은 식(5)에 따라 각각의 8×8 DCT 블록의 4×4 저주파성분을 변형한 후에 4×4 IDCT를 사용하여 크기를 줄인다. 그리고 나서, 공간영역에서 크기가 줄어든 영상은 각각의 8×8 DCT 블록으로 부호화된다. DCT 영역에서의 영상 축소 알고리즘은 다음과 같이 설명된다.

입력 영상의 각각의 8×8 DCT 블록에 대해서

1단계: B 에서의 DCT 계수를 4×4 DCT 블록 \hat{B} 로 다음과 같이 변경한다.

$$\hat{B}(u, v) = \frac{1}{2} B(u, v) \cos\left(\frac{u\pi}{16}\right) \cos\left(\frac{v\pi}{16}\right) \quad (6)$$

$$u, v=0, 1, 2, 3$$

2단계: $\hat{B}(u, v)$ 에 4×4 IDCT를 적용한다.

3단계: 2단계에서 크기가 줄어든 영상에 8×8 DCT를 수행한다.

[1]에서 유도된 최소 행렬을 사용하여 2단계와 3단계를 하나의 단계로 결합함으로써 효율적인 계산이 가능하다. 2단계에서 4×4 IDCT를 할 때, [1]에서의 계산량은 원래영상의 한 픽셀당 1.25번의 곱셈과 1.25번의 덧셈이 필요하다. 본 논문에서 제시된 방법은 DCT 계수를 변경하기 위해 $256(16 \times 16$ 블록) 픽셀에 대해 60번의 부가적인 곱셈이 필요하기 때문에 원래영상의 한 픽셀당 1.48번의 곱셈과 1.25번의 덧셈이 필요하다.

2. 영상의 확대 방법

DCT 영역에서 영상의 확대는 원 영상의 8×8 DCT 계수를 변경하고 이를 16×16 IDCT를 하여 영상을 확대한다. 그리고 나서, 공간영역에서 확대된 영상을 각각 8×8 DCT 블록으로 부호화한다. DCT 영역에서의 영상 확대 알고리즘은 다음과 같이 설명된다.

입력 영상의 각각의 8×8 DCT 블록에 대해서

1단계: 8×8 DCT 블록 B 를 16×16 DCT 블록 \bar{B} 로 다음

과 같이 변경한다.

$$\bar{B}(u, v) = \begin{cases} \frac{2B(u, v)}{\cos\left(\frac{u\pi}{16}\right)\cos\left(\frac{v\pi}{16}\right)} & u, v = 0, 1, \dots, 7 \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

2단계: $\bar{B}(u, v)$ 에 16×16 IDCT를 적용한다.

3단계: 2단계에서 크기가 늘어난 영상에 8×8 DCT를 수행한다.

본 논문에서 제안된 영상 확대 방법은 [2]에서의 최소 행렬과 비슷하게 2단계와 3단계를 결합해서 하나의 단계로 효과적으로 계산될 수 있다. B'를 식(7)을 사용하여 변경된 8×8 DCT 블록이라고 하자. 그러면 확대된 16×16 DCT 블록 \bar{B} 는 다음과 같다

$$\bar{B} = \begin{bmatrix} B' & 0 \\ 0 & 0 \end{bmatrix} \quad (8)$$

T_8 과 T_{16} 를 8 DCT 행렬과 16 DCT 행렬이라고 하면, 4개의 이웃하는 8×8 블록들 (b_1, b_2, b_3, b_4)로 구성된 16×16 블록은 다음과 같이 계산된다.

$$\begin{bmatrix} b_1 & b_2 \\ b_3 & b_4 \end{bmatrix} = T_{16}' \bar{B} T_{16} = \begin{bmatrix} T_{16LL}' B' T_{16LL} & T_{16LL}' B' T_{16LH} \\ T_{16LH}' B' T_{16LL} & T_{16LH}' B' T_{16LH} \end{bmatrix} \quad (9)$$

그러므로 식(9)로부터 확대된 영상의 4개의 DCT 블록들 (B_1, B_2, B_3, B_4)은 다음과 같이 얻어질 수 있다.

$$\begin{bmatrix} B_1 & B_2 \\ B_3 & B_4 \end{bmatrix} = \begin{bmatrix} T_8 T_{16LL}' \bar{B} T_{16LL} T_8' & T_8 T_{16LL}' \bar{B} T_{16LH} T_8' \\ T_8 T_{16LH}' \bar{B} T_{16LL} T_8' & T_8 T_{16LH}' \bar{B} T_{16LH} T_8' \end{bmatrix} \quad (10)$$

식(10)에서 $T_{16LL} T_8'$ 과 $T_{16LH} T_8'$ 은 행렬 요소의 부호를 제외하곤 동일한 행렬이고 요소들의 반은 0을 가진다.

$T_{16LL} T_8'$ 과 $T_{16LH} T_8'$ 을 $(A+B)$ 와 $(A-B)$ 의 형태로 분해함으로써 식(10)은 다음과 같이 다시 쓰여질 수 있다.

$$\begin{aligned} B_1 &= (A+B)' \bar{B} (A+B) \\ B_2 &= (A+B)' \bar{B} (A-B) \\ B_3 &= (A-B)' \bar{B} (A+B) \\ B_4 &= (A-B)' \bar{B} (A-B) \end{aligned} \quad (11)$$

$$A = \begin{bmatrix} .7071 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & .2986 & 0 & .0241 & 0 & .0071 & 0 & .0018 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & .5446 & 0 & -.0951 & 0 & -.0235 & 0 & -.0057 \\ 0 & 0 & .7071 & 0 & 0 & 0 & 0 & 0 \\ 0 & -.2219 & 0 & .4008 & 0 & .0493 & 0 & .0110 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & .1509 & 0 & .4971 & 0 & -.1078 & 0 & -.0196 \end{bmatrix}$$

$$B = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ .6376 & 0 & -.0585 & 0 & -.0125 & 0 & -.0039 & 0 \\ 0 & .7071 & 0 & 0 & 0 & 0 & 0 & 0 \\ -.2153 & 0 & .3812 & 0 & .0436 & 0 & .0128 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ .1326 & 0 & .5081 & 0 & -.1061 & 0 & -.0253 & 0 \\ 0 & 0 & 0 & .7071 & 0 & 0 & 0 & 0 \\ -.0985 & 0 & -.2024 & 0 & .4065 & 0 & .0476 & 0 \end{bmatrix}$$

식(11)을 효율적으로 계산하면 제안된 영상 확대 방법을

이용해서 확대된 영상의 한 픽셀당 3.62번의 곱셈과 4.25번의 덧셈으로 DCT 영역에서의 영상 확대가 가능하다.

IV. 실험 및 결과

본 논문에서는 [1]과 [2]에서 사용된 것과 같은 JPEG으로 부호화된 Lena 영상, Watch 영상, F-16 영상 그리고 Cap 영상을 사용하였다. 제안된 방법으로 얻어진 Lena 영상의 축소 영상과 축소 후 확대된 영상을 그림 1에 나타내었다. 축소한 후 확대된 영상은 원 영상과 거의 동일함을 알 수 있다.



(a) Lena 원 영상



(b) 제안된 방법으로 축소 후 확대된 영상



(c) 제안된 방법에 의해 축소된 영상
그림 1. 축소 후 확대된 영상 및 축소 영상

표 1은 각각의 방법에 따른 PSNR값을 나타낸다. 제안된 방법으로 영상을 축소한 후 확대한 영상의 화질(PSNR)이 가장 좋음을 알 수 있다.

표 1. 축소 후 확대한 영상의 PSNR값

	PSNR값			
	DA	SB	TR	Proposed
Lena	34.66	34.81	34.94	34.99
Watch	29.29	29.60	29.75	29.84
Cap	34.37	34.37	34.41	34.41
F-16	32.46	32.72	32.83	32.89

V. 결론

본 논문에서는 DCT 영역에서 직접 영상을 축소하고 확대하기 위한 방법을 제안하였다. 영상을 축소하기 위한 변경된 IDCT 방법이 제안되었으며, 이에 상응하여 영상 확대에 사용되는 방법이 동시에 제안되었다. 본 논문에서 제안된 방법을 사용하여 DCT 영역에서 영상을 축소한 후 확대한 결과 기존의 다른 방법에 비하여 높은 PSNR값을 가졌다.

참고 문헌

- [1] R. Dugad and N. Ahuja, "A Fast Scheme for Image Size Change in the Compressed Domain," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 11, no. 4, pp. 461-474, Apr. 2001.
- [2] J. Mukherjee and S. K. Mitra, "Image Resizing in the Compressed Domain Using Subband DCT," *IEEE Trans. Circuit and Systems for Video Technology*, vol. 12, no. 7, pp. 620-627, Jul. 2002.
- [3] S. H. Jung, S. K. Mitra and D. Mukherjee, "Subband DCT: Definition, analysis and applications," *IEEE Trans.*