

# DCT 영역 영상 크기 조절 방법들에 대한 PSNR 비교

정회원 김도년\*, 최윤식\*\*

## PSNR Comparison of DCT-domain Image Resizing Methods

Donyeon Kim\* and Yoonsik Choe\*\*, *Regular Members*

### 요 약

비디오 프레임의 크기를 축소하거나 확대할 때, 응용에 따라서는 입력 및 출력이  $8 \times 8$  블록 DCT 계수들로 구성되도록 할 필요가 있다 선형 변환이고 유니터리(unitary) 변환의 일종인 DCT에는 행렬 곱셈에 대한 분배 법칙이 성립한다 이러한 사실을 이용하여 두가드, मुखर्जी, 박 등은 DCT 영역에서 비디오 프레임들의 크기를 축소하는 방법들을 제안하였다 이러한 방식으로 영상을 축소 후 확대하면 원 영상의 저 주파수 DCT 계수들이 잘 보존된다 즉, 원 영상(축소되기 전의 영상)과 예측된 영상(축소 후 확대된 영상)의 차이를 부호화 해야 되는 경우 부호화 효율이 매우 높아진다 이러한 것은 스케일러빌리티를 이용한 비디오 부호화에 비할만한 사실이다 본 논문에서는 이전의 방식들의 연장선 상에서 가로 세로 각각 2:1로 축소하고 다시 2:1로 확대할 때 DCT 블록의 크기를 다양히 하였다 실험에 의하면 DCT 블록 크기를 크게 할수록 PSNR 값이 커짐을 알 수 있었다 그러나, 계산상의 복잡도 역시 커질 것으로 예상된다 본 논문의 실험 결과는 압축 영역 영상 축소 및 확대를 위한 고속 알고리즘 개발에 중요한 데이터가 될 것으로 생각한다

Key Words DCT, Image/video downsizing/upsizing, MPEG, JPEG, Scalability

### ABSTRACT

Given a video frame in terms of its  $8 \times 8$  block-DCT coefficients, we wish to obtain a downsized or upsized version of this frame also in terms of  $8 \times 8$  block DCT coefficients The DCT being a linear unitary transform is distributive over matrix multiplication This fact has been used for downsampling video frames in the DCT domains in Dugad's, Mukherjee's, and Park's methods. The downsampling and upsampling schemes combined together preserve all the low-frequency DCT coefficients of the original image This implies tremendous savings for coding the difference between the original frame (unsampled image) and its prediction (the upsampled image) This is desirable for many applications based on scalable encoding of video. In this paper, we extend the earlier works to various DCT sizes, when we downsample and then upsample of an image by a factor of two Through experiment, we could improve the PSNR values whenever we increase the DCT block size However, because the complexity will be also increase, we can say there is a tradeoff The experiment result would provide important data for developing fast algorithms of compressed-domain image/video resizing

### 1. 서 론

MPEG, H.261, H.263 등 대부분의 동영상 압축

방식들이 DCT(Discrete Cosine Transform, 이산역 현변환)와 같은 변환을 이용한다 또한 영상 편집, 특수 효과 등을 구현하기 위하여 실시간 처리가 필

\* 연세대학교 IT 연구단

\*\* 연세대학교 전기전자공학부

논문번호 KICS2004-09-190, 접수일자 2004년 9월 9일

요하다. 특수 효과에는 영상크기 조절하기, 밝기 조정, 필터링, 마스킹, 영상의 회전등이 있다. 이 경우 영상 압축을 풀어서 영상 처리를 하고 다시 압축하게 되면 계산상의 복잡도가 커지는 문제점이 있다 그래서 워크스테이션 등에서 이 같은 처리를 하는 것은 큰 부담이 될 수 있다 이러한 이유로 직접 변환 영역에서 고속알고리즘을 개발하여 압축을 풀지 않거나 최소한 계산상의 병목현상을 제거하는 연구가 최근에 활발히 진행 되고 있다[1]-[7].

예를 들어 다자간 영상 회의를 진행한다고 하자. 이때 참여자들은 각자의 화면상 내에 여러 개의 창들을 통해서 참여자를 모두 볼 수 있다 참여자들은 창들의 크기를 마음대로 조절하고자 할 것이다 각각의 워크스테이션은 하나의 비디오 스트림만을 처리할 수 있으므로, 서버는 전체 참여자들로부터 온 스트림들을 참가자의 요구를 반영하여 하나의 스트림으로 재구성 하여야 한다 만약 한 참가자가 자기 화면에 다른 참가자의 창을 가로 및 세로 각각 2.1로 줄이고자 하면 서버에서 영상을 축소하고 디지털 비디오 스트림을 재구성해야 된다[1][6].

기존의 방식은 서버에서 압축된 비디오 스트림을 완전히 풀고, 필요한 연산을 수행한 후 다시 압축하여 원하는 참가자에게 비디오 스트림을 보내는 것이다. 이때 DCT와 IDCT 연산으로 인한 계산상의 부하가 매우 크기 때문에 DCT 영역에서 이러한 처리를 할 수 있는 고속알고리즘이 필요하다

본 논문에서는 이러한 방법들을 정리하고 실험을 통하여 이 방법들의 화질을 비교함으로써 DCT 영역 영상 크기 조절에 관련된 일련의 연구에 중요한 데이터를 제공 한다.

앞에서 '회상 회의 시나리오'를 설명했지만 다음과 같이 인터넷에 응용하는 경우도 생각해 볼 수 있다 거리상으로 멀리 떨어진 이미지 혹은 영상 데이터 베이스를 검색할 때, 처음에는 축소된 영상을 클라이언트에게 보내고 크라이언트가 요청하거나 클라이언트가 관심이 있는 영상만 원영상 크기로 크게 만들어 보낸다 이때 공간영역 (화소 영역)에서 직접 여러가지 내삽법 기술들을 이용하여 영상 크기를 조절할 수 있다. 그러나 효율적인 이미지 저장을 위해 영상들은 일반적으로 변환 영역의 데이터로 압축되어 있으므로 압축된 스트림 상에서 이미지 크기를 조절하는 알고리즘이 필요하다

공간 스케일링 필터에서는 공간해상도가 낮은 베이스 층과 공간 해상도가 높은 인텐스먼트 층 혹은 높은 층으로 나누어 동영상을 압축한다. 인텐스

먼트 층은 베이스 층을 이용하여 공간 예측 부호화한다. 그러므로, 낮은 해상도의 동영상이 필요하다면 베이스 층만 압축을 풀면 된다. 그러나 해상도가 높은 동영상이 필요하다면 베이스 층과 인텐스먼트 층 모두의 압축을 풀고 두 동영상을 결합한다 공간 스케일링 필터는 HDTV와 표준 해상도 TV와 호환성을 유지시키는데 이용된다 비슷한 예로, 우선성이 이중인 넷워킹상에서 동영상을 전송할 때 우선도가 높은 채널에는 낮은 해상도의 동영상을 보내고, 우선도가 낮은 채널에는 인텐스먼트 층의 데이터를 보낸다. 본 논문에서 이와 같은 응용을 '공간 스케일링 필터 시나리오'라고 부른다[3]-[5] [7].

본 논문의 2장에서 기존의 DCT 영역 영상 조절 알고리즘에 대해서 알아보고 그 방법들의 확장 방안에 대해서 알아보고, 그 방법들을 확장했을 때 PSNR 성능이 어떻게 되는지 3장에서 실험을 통하여 살펴보고 4장에서 간단히 결론을 맺는다.

## II. DCT 영역 영상 크기 조절 알고리즘들

영상을 축소할 때, 먼저 저역 통과 필터링을 하고 다운샘플링을 하게 된다 이때 필터의 특성이 줄어든 영상의 화질을 결정한다. DCT 영역 영상 크기 조절 알고리즘은 크게 나누어 네 가지로 볼 수 있다

먼저, 멀하브(Merhav)는 DCT 영역에서 영상의 크기를 줄일 때 저역 통과 필터로 쌍일차 필터를 이용하였다 계산 량을 줄이기 위하여 DCT와 양자화를 결합하였다 이때 이용한 고속 DCT는 곱셈이 밖으로 몰려 있어서 양자화에 흡수되기 쉬운 구조를 가지고 있는 고속 위노그라드(Winograd) DCT를 이용하였다. 멀하브의 방법은 '회상 회의 시나리오'에 적용하기 위하여 개발하였다. 또한 영상을 확대 하는 방법에 대하여는 언급하지 않았다 [1]

멀하브의 방법 이후 두가드(Dugad) 방법이 제안되었다 앞에서 설명하였듯이, 영상을 축소 할 때 먼저 저역 통과 필터링을 하고 다운샘플링을 한다 멀하브의 방법에서 저역 통과 필터는 DCT와 무관하게 선택된다 두가드는 필터 행렬의 계산 량을 줄이는 것보다 필터 행렬을 DCT한 것의 계산 량이 적도록 저역 통과 필터를 설계하였다. 즉 저역통과 필터와 다운샘플링을 결합하여 직접 DCT 영역에서 처리하였다[3]

구체적으로 설명하면 다음과 같다. 그림 1 (a)에 나타난 것과 같이 영상을 8×8 크기 블록으로 분할

하고, 각 블록을  $8 \times 8$  DCT를 한다(그림 1 (a) 및 (b)에서  $N=8$ ). 크기가  $8 \times 8$ 인 DCT 계수에서 저 주파수 쪽의 계수  $4 \times 4$  만 떼어내서  $4 \times 4$  IDCT를 취하면 바로 저역 통과되고 다운샘플링된 데이터를 얻을 수 있다. 크기가 줄어든 영상은 원영상의 모든 저 주파수 성분들을 그대로 가지게 되는 장점이 있다. 영상 크기를 확대할 때는 다음과 같다. 영상을  $4 \times 4$  크기 블록으로 분할하고, 각 블록을  $4 \times 4$  DCT를 한다. 이 블록을 저주파수 DCT 계수로 하고, 나머지는 영으로 채운 크기가  $8 \times 8$ 인 DCT 블록을 만든 후, 각 블록을  $8 \times 8$  역 DCT를 취해주면, 바로 업샘플링되고 저역 통과된 데이터, 즉 확대된 영상을 얻을 수 있다.

이전의 멀허브 방법은 영상 크기를 줄이는 방법에 상응하는 영상 확대 방법에 대한 대책이 없다. '공간 스케일러빌리티 시나리오'에선 영상 확대방법이 꼭 필요하다. 즉, 베이스 층에는 저 해상도 영상이 전송되고 이 영상을 확대하여 고 해상도 영상을 예측한다. 두가드는 영상 축소와 같은 방법으로 영상을 확대하므로 원영상의 저 주파수 계수들이 모두 보존 된다. 따라서 인헨스먼트 층에서는 원 영상과 예측된 영상의 차이를 전송하므로 비트 량을 상당히 줄일 수 있다. 즉, 인헨스먼트 층에는 고 주파수 계수들만 전송하면 된다. 그렇지만, 영상 축소 방법은 어느 것을 쓰든 상관치 않고 영상 확대 방법만 한가지로 고정하여 사용할 수 있다. 실험에 의하면 영상 확대방법이 주관적 화질 및 PSNR을 결정한다. 이러한 특징은 영상의 줄이기 방법이 이미 결정되어 있고, 영상의 확대 방법만 선택할 수 있는 경우에 매우 중요한 사실이다.

두가드의 방법을 개선한 것이 묵허지(Mukherjee) 방법이다. 두가드는 영상크기를 크게할 때  $8 \times 8$  DCT 계수를 IDCT한 후 네개의  $4 \times 4$  블록으로 만든 후 각 블록에 대하여  $4 \times 4$  DCT를 취한다. 이후 각  $4 \times 4$  DCT 계수를  $8 \times 8$  DCT 계수 블록의  $4 \times 4$  저주파수 계수로 사용하고 그 이외의 부분을 영으로 채운 후  $8 \times 8$  IDCT를 함으로써 원 영상 크기보다 가로 세로 각각 두 배씩 큰 영상을 얻는다. 묵허지는 여기서 더 발전하여 영상을 확대할 때  $8 \times 8$  DCT 계수를  $16 \times 16$  DCT 계수 블록의  $8 \times 8$  저 주파수 계수로 사용하고 그 이외의 부분을 영으로 채운 후  $16 \times 16$  IDCT를 함으로써 원 영상 크기보다 가로 세로 각각 2배씩 큰 영상을 얻는다(그림 1(a)에서  $N=8$ , (b)에서  $N=16$ ). 이렇게 함으로써 고속 알고리즘을 만들었을 때 계산 량이 늘어나지만, 화질이 좋아지는 효과가 있다[4].

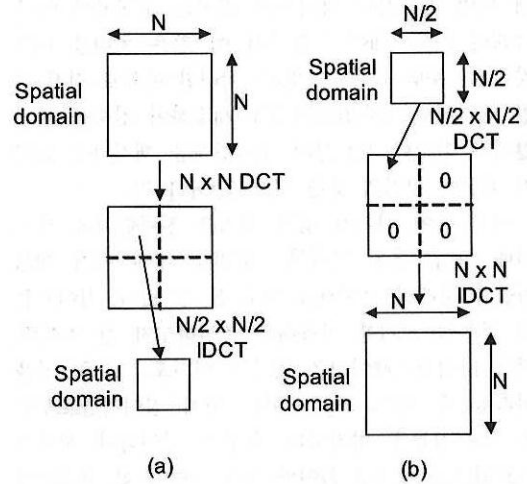


그림 1. (a) DCT를 이용한 영상 축소 (b) DCT를 이용한 영상 확대

묵허지의 방법을 개선한 것이 박의 방법이다. 영상을 크게 할 때는 묵허지의 방법과 동일하지만 영상의 크기를 작게 할 때 묵허지의 방법을 발전시켰다. 즉, 개념적으로 설명하면 다음과 같다. 원 영상을  $16 \times 16$  블록들로 구분한 후  $16 \times 16$  DCT처리를 한다. 크기가  $16 \times 16$ 인 DCT 계수에서 저주파수 쪽의 계수  $8 \times 8$  만 떼어내서  $8 \times 8$  IDCT를 취하면 바로 저역 통과되고 다운샘플링된 데이터를 얻을 수 있다(그림 1 (a) 및(b)에서  $N=16$ ) [5].

지금까지의 방법을 정리하면 다음과 같다. 멀허브는 쌍 일차 외삽법을 이용하여 영상을 축소하였다. 두가드는  $8 \times 8$  DCT 계수를  $4 \times 4$  DCT 계수로 만드는 방법으로 영상을 축소하고,  $4 \times 4$  DCT 계수를 이용하여  $8 \times 8$  DCT를 만드는 방법으로 영상을 확대하였다. 묵허지의 방법과 박의 방법은 두가드의 방법과 유사하나 DCT 블록의 크기가 다름을 알 수 있다.

### III. 실험을 통한 영상화질 비교

앞장에서 소개한 방법들의 화질을 비교하기 위하여 실험을 하였다. 여기서,  $8 \times 8$  DCT 계수의 저 주파수  $4 \times 4$ 만 골라내어  $4 \times 4$  IDCT를 함으로써 영상 크기를 줄이는 방법을 8:4 다운샘플링이라고



부르고(그림1 (a)에서 N=8), 4×4 DCT 를 8×8 DCT 블록의 저주파수 계수로 이용하고 나머지를 영으로 채운 뒤 8×8 IDCT를 함으로써 영상을 확대하는 방법을 4:8 업샘플링이라고(그림 1 (b)에서 N=8) 이름 하였다. 이와 같은 방식으로, 다음과 같이 네 가지 방법에 대하여 실험을 하였다.

방법1: 16:8 다운샘플링과 4:8 업샘플링(그림1 (a)에서 N=16, (b)에서 N=8)

방법2(두가드 방법): 8:4다운샘플링 후 4:8 업샘플링(그림1 (a)에서 N=8, (b)에서 N=8)

방법3(묵허지 방법): 8:4다운샘플링 후 8:16 업샘플링(그림1 (a)에서 N=8, (b)에서 N=16)

방법4(박의 방법): 16:8다운샘플링 후 8:16 업샘플링(그림1 (a)에서 N=16, (b)에서 N=16)

주의할 점은 네 방법 모두, 가로 세로 각각 다운 및 업샘플링율이 2:1로 동일하다. 표1로부터 방법2보다는 방법3이, 방법3보다는 방법4의 객관적 화질이 더 좋음을 알 수 있다. 특이한 것은 방법1의 화질이 가장 좋지 않다는 것이다. 즉 방법2보다 더 나은 좋음을 알 수 있다. 따라서 업샘플링(영상 확대)시에 방법을 잘 선택해야 함을 알 수 있다. 또한 화질이 좋은 방법일수록 계산 시에 계산 량이 많아지는 점도 고려해야 한다.

표 1. 영상 크기를 가로 세로 각각 1/2로 축소한 후 같은 비율로 확대했을 때의 PSNR 비교

실험 영상	PSNR (dB)			
	방법1	방법2	방법3	방법4
Lena	34.88	35.17	32.50	33.85
Watch	28.94	29.23	29.68	30.04
Cap	34.07	34.37	34.39	34.70
F-16	32.13	32.45	32.82	33.18

주관적 평가를 위한 그림 2에서 (b)의 쌍 일차 내삽법을 이용한 방법이 가장 흐릿함을 알 수 있다. 두가드 방법인 (c)도 우수하지만, (d)의 박의 방법이 주관적 화질에서도 가장 좋다. 즉, 모자의 털 부분과 눈 부분이 또렷함을 알 수 있다.

실험을 확장하여, 영상의 크기를 가로 세로 각각 두 배로 확대할 때 DCT 크기를 계속 크게 하였을 경우에 대한 실험 결과를 그림 3에 나타내었다. 그림 3으로부터 이제까지 설명한 방법들의 연장선상에서 DCT 크기를 크게 할수록 PSNR 값이 커짐을 알 수 있다. 그러나 PSNR값의 증가율은 둔화됨을

알 수 있다. 현재 고속 알고리즘은 M=0 (두가드 방법), M=1 (박의 방법)에 대하여 개발되었다. 이보다 DCT 크기를 크게 (M=2,3,4) 하여 높은 PSNR 값을 가지는 방법을 개발할 수 있지만, 그에 따라 계산 량도 많이 늘어나는 문제점이 있다.

실제 JPEG 부호화기를 이용하여 압축률을 바꾸어 가며 앞과 동일한 실험을 행하였을 때, 앞과 같은 결과를 얻을 수 있었다(그림 4). 압축률이 높아지면 전체적인 PSNR 값이 떨어지기 때문에 DCT 블록의 크기를 크게 한 방법들과 작게 한 방법과의 차이가 줄어든다.

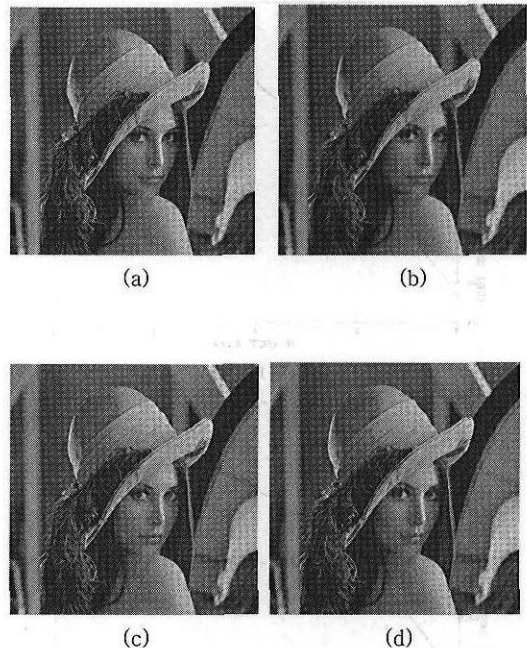
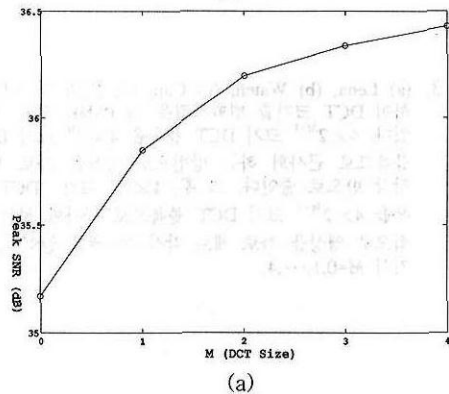
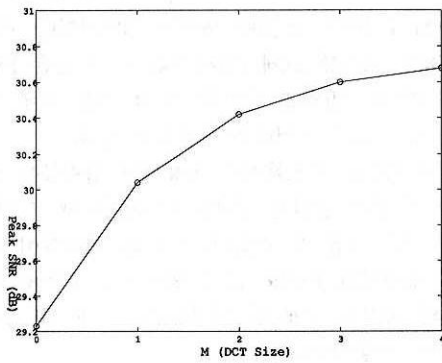
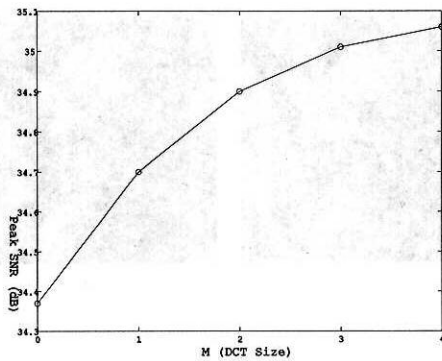


그림 2. (a) Lena의 원영상. (b) 쌍일차 내삽법(멀허브 방식), (c) 두가드 방식, (d) 박의 방식으로 Lena 영상 크기를 가로 세로 각각 1/2로 축소 후 같은 비율로 확대 했을 경우의 Lena 영상들.

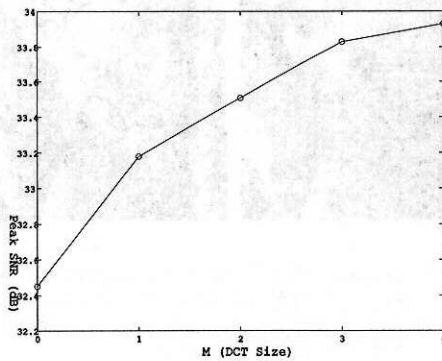




(b)

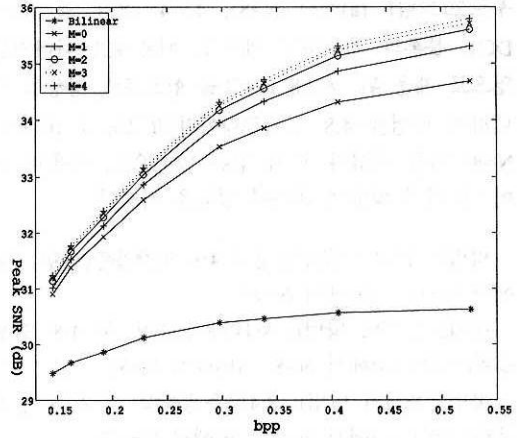


(c)

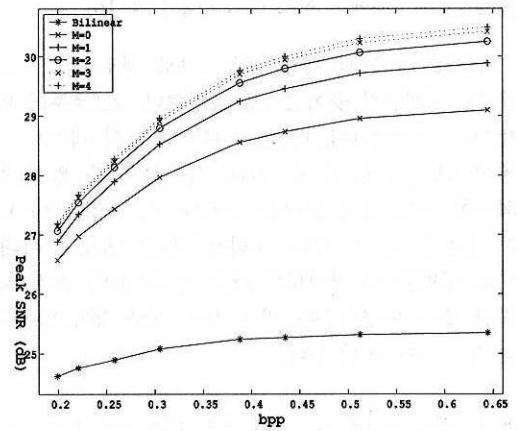


(d)

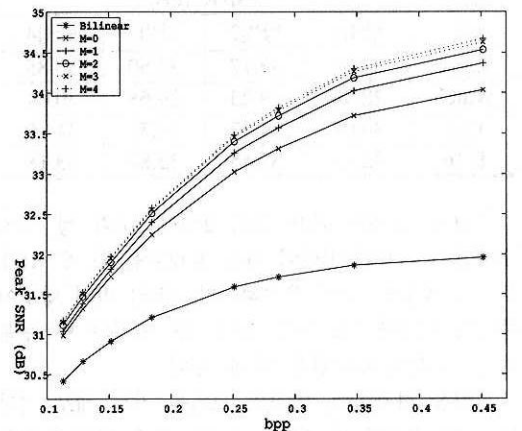
그림 3. (a) Lena, (b) Watch, (c) Cap, (d) F-16 영상에 대하여 DCT 크기를 변화시켰을 때 PSNR 값을 구하였다.  $4 \times 2^{M+1}$  크기 DCT 블록을  $4 \times 2^M$  크기 DCT 블록으로 근사화 하는 방법으로 영상을 가로 세로 각각 반으로 줄인다. 그 후,  $4 \times 2^M$  크기 DCT 블록을  $4 \times 2^{M+1}$  크기 DCT 블록으로 근사화 하는 방법으로 영상을 가로 세로 각각 두 배로 늘인다. 여기서  $M=0,1,\dots,4$ .



(a)



(b)



(c)



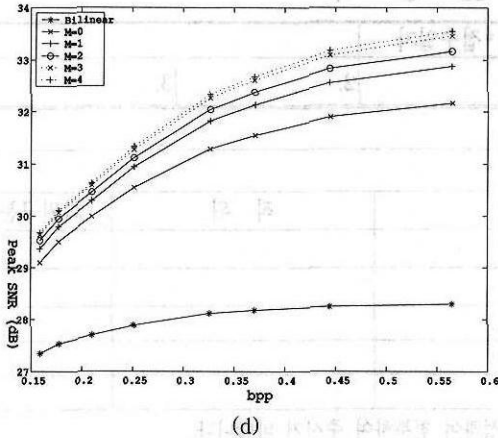


그림 4. JPEG 부호화기의 입력으로 (a) Lena, (b) Watch, (c) Cap, (d) F-16 영상을 이용하였다. JPEG으로 압축된 상태에서 DCT 크기를 변화시켰을 때, 그리고 DCT 크기를 고정시키고 압축률을 변화시켰을 때 PSNR 값을 구하였다.  $4 \times 2^{M+1}$  크기 DCT 블록을  $4 \times 2^M$  크기 DCT 블록으로 근사화 하는 방법으로 영상을 가로 세로 각각 반으로 줄인다. 그 후,  $4 \times 2^M$  크기 DCT 블록을  $4 \times 2^{M+1}$  크기 DCT 블록으로 근사화 하는 방법으로 영상을 가로 세로 각각 두 배로 늘린다. 여기서  $M=0,1,\dots,4$ .

#### IV. 결론

최근 활발히 연구되고 있는 DCT 영역 영상 축소 및 확대에 대한 방법을 살펴보았다. 최근의 두가드, 목허지, 박의 방법들은 DCT의 저 주파수만 이용하여 영상을 축소 및 확대하였다. 본 논문에서는 이 방법들에서 사용하는 DCT 크기가 클수록 화질이 좋아짐을 실험을 통하여 확인하였다.

즉, Lena 영상의 경우에, 사용하는 DCT 크기가 가로 세로 각각 배가 될 때 마다 0.68 dB, 0.35 dB, 0.14 dB, 0.09 dB 향상 되었다(그림 3 (a)). 이때 다운 및 업 샘플링 율은 동일하게 하고 비교 하였다.

앞으로 이 방법들에 대한 고속알고리즘 개발에 본 논문의 실험 결과가 중요한 데이터를 제공할 것으로 생각한다.

#### 참 고 문 헌

[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, no. 3, pp. 468-476, June, 1997.

[2] Q. Hu and S. Panchanathan, "Image/video spatial scalability in compressed domain," *IEEE Trans. Ind. Electron.*, vol. 45, pp. 23-31, Feb., 1998.

[3] R. Dugad and N. Ahuja, "A fast scheme for image size change in the compressed domain," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 11, no. 4, pp. 461-474, Apr., 2001.

[4] J. Mukherjee and S. K. Mitra, "Image resizing in the compressed domain using subband DCT," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 12, no. 7, pp. 620-627, July, 2002.

[5] H. Park, Y. Park, and S. Oh, "L/M-fold image resizing in block-DCT domain using symmetric convolution," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 12, no. 9, pp. 1016-1034, Sept., 2003.

[6] D. Kim and Y. Choe, "Fast algorithm for discrete cosine transform (DCT)-domain image downsampling using Winograd DCTs," *Optical Engineering*, vol. 42, no. 9, pp.2485-2486, Sept., 2003.

[7] R. Dugad and N. Ahuja, "A scheme for spatial scalability using non-scalable encoders," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 10, pp. 993-999, Oct., 2003.

김도년(Donyeon Kim)	정희원
현재 : 연세대학교 IT 연구단	
<관심분야> 영상 신호 처리, 영상 압축	
최윤식(Yoonsik Choe)	정희원
현재 : 연세대학교 공과대학 전기전자공학부	
<관심분야> 영상 신호 처리, 영상 압축	