

프레임 차분 체적양자화에 의한 동영상의 복원특성

최재윤*, 임낙현*, 김태효**

*경남대학교 전자공학과, **경남대학교 정보통신공학부

The Reconstruction Characteristics of Moving Images with Frame Difference Volume Quantizer

Choi Jae-yoon*, Leem Nack-hyun*, Kim Tae-hyo**

*Department of Electronics Engineering in Kyungnam University

**Division of Information and Communication in Kyungnam University

요약문

본 논문에서는 영상 시퀀스의 이전 프레임과 현재 프레임 사이의 차분 데이터로 이루어진 차분 프레임에 대하여 제안한 FDVQ(프레임 차분 체적양자화)를 적용했을 때, 복원된 영상의 원 영상에 대한 해상도(PSNR)를 분석하였다.

시험 동영상(moving image)으로 실험한 결과, 저주파 성분을 많이 포함하는 차분 영상 데이터에 대하여 높은 압축률에서 제안된 FDVQ 방법으로 얻은 복원 해상도는 2D-DCT와 VQ를 이용한 방법보다 개선된 결과를 나타내었다.

PSNR 결과에 대한 영상 특성의 분석 결과에서 FDVQ를 이용한 방법은 VQ 방법보다 압축률은 비슷하였으나 해상도는 평균 0.39dB 정도 향상되었다. 또한, 평균압축률 약 0.05bpp에서 2D-DCT 방법에 비해 고주파 성분을 많이 포함하는 경우는 0.8dB 이상 향상되었으며, 저주파 성분을 많이 포함하는 경우는 평균 0.4dB 정도 향상된 복원영상(reconstructed image)을 얻을 수 있었다.

I. 서론

멀티미디어 분야는 인터넷, 정보통신 등으로 통신, 방송, 저장매체 등의 이용에 주도적 역할을 해 오고 있으며, 정보통신을 이용한 실시간(real time)기술의 도입으로 동영상까지 지원되고 있으

나 막대한 정보량 때문에 실시간을 이용한 동영상 통신은 아직도 영상 신호를 처리하는 압축 기술이 필수적으로 따르게 된다. 영상 신호를 압축 없이 통신에 이용할 경우, NTSC 칼라 비디오 신호는 30 frame/sec으로 디지털 전송시 요구되는 전송량은 약 248Mbps이며, HDTV 칼라 신호인 경우는 채널 용량(channel capacity)이 1.5Gbps나 된다.

이러한 막대한 양의 영상정보는 정보의 공간적 분포가 전 영상부분에 나타나므로 정보를 분석하여 처리하기가 매우 어렵게 된다. 따라서 정보를 변화되는 크기(주파수 성분 등)에 따라 일정 부분으로 집중되는 주파수 영역별로 나타내어 주파수 변환 방법으로 분석하기가 편리하게 된다[1-3].

동영상 압축기술은 근래에 와서 이산 웨이블릿 변환(discrete wavelet transform)[4], 부밴드 부호화(subband coding)[5] 등에 기초를 두고 디지털 신호처리 분야에서 제안되어졌지만, DCT는 여전히 영상 및 동영상 압축에 대하여 중요한 변환 부호화 기술로 남아있다.

그러나, 현재 영상 압축 표준으로 사용되고 있는 2D-DCT는 동영상의 시간축에 대한 적용이 곤란하고 압축 효과가 실용적이지 못하여 영상의 실시간 처리에 적합하지 않고 MPEG 비디오 코딩[6]에서의 시간소비와 움직임 추정에 의한 손실(lossy) 동작 등의 결함이 생긴다.

동영상 시퀀스에 대해 움직임 추정은 전체를 조사하지 않는 한 좋은 움직임 벡터를 얻지 못하는

데, 2차원 DCT[7]에서는 약간의 손실을 가져오더라도 양자화를 이용한 압축률을 높이는 방법과 좋은 해상도를 얻기 위해서 압축률을 낮추는 방법을 이용해야 한다.

연속 영상 시퀀스에 대해서 시간축을 고려한 3차원으로 해석할 수 있다. 2차원을 확장한 3D-DCT는 대칭적 코덱과 동질적(homogeneous) 비디오에 대한 높은 압축률과 같은 고유의 장점[8]을 가지고 있다.

본 논문에서는 연속적인 동영상 시퀀스의 프레임간 차분 성분을 3D-DCT 계수로 분석하여 계수 특성에 맞는 프레임 차분 체적 양자화기(frame difference volume quantizer : FDVQ)를 이용하여 여러 가지 영상의 주파수 성분에 대한 복원 영상의 해상도를 비교 분석하고자 하였다.

II. 본 론

2.1 이론적 배경

DCT는 영상 데이터 압축 기법의 하나로 MPEG, H.261 등의 영상 압축 표준안에서 기반으로 하고 있다. 이 변환은 블록기반 변환 방식으로 적은 수의 저주파수 신호 영역에 영상 정보를 집중시키고 나머지 영상 정보들은 '0' 또는 '0' 부근의 작은 값을 갖도록 처리함으로써 많은 양의 영상 데이터를 적은 정보량으로 만들어 전송하고자 하는 데이터를 압축 처리할 수 있다.

즉, 영상신호는 이웃하는 픽셀과는 차이가 거의 없고, 픽셀간의 상관성이 매우 높아 공간 주파수 관점에서 저주파 성분에 해당하기 때문에 저주파 성분에 대하여 정보 신호의 의미를 두게된다.

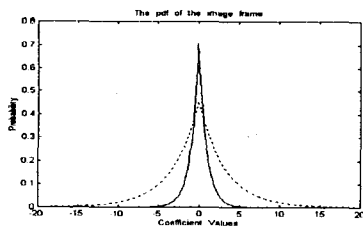


그림 2.1 AC 계수 확률밀도함수 분포곡선 비교
Fig. 2.1 The pdf distribution diagram comparison of AC Coefficients

그림 2.1은 기존의 3D-DCT에 의한 것과 제안한 FDVQ의 차분 영상에 대한 3D-DCT의 AC 계수에 대한 확률밀도함수 관계를 나타내었다.

접선으로 나타난 분포 곡선은 기존의 3D-DCT에 의한 AC 계수의 분포 함수이며 실선으로 나타난 것은 제안한 방법에 의한 분포 함수이다.

2.2 3D-DCT 특성

연속되는 영상에 대하여 순간적인 2차원 평면공간에서 하나의 시간함수를 나타내는 변수를 추가하게 되면, $N_R \times N_C$ 블록은 $N_F \times N_R \times N_C$ 로 표현되는 큐빅블록(cubic block)으로 확장할 수 있다. N_F 는 3D-DCT 영역에서 시간축으로 큐빅블록을 고려한 $N_R \times N_C$ 블록의 연속적인 프레임 수를 의미한다.

2D-DCT에 대한 표현식에 시간함수를 추가하면, 일반적인 3D-DCT와 3D-IDCT의 표현식은 식(2.1)과 식(2.2)와 같이 정의할 수 있다[9][10].

- 3D-DCT :

$$S(w, v, u) = a_{3D}(w, v, u) \sum_{z=0}^{N_F-1} \sum_{y=0}^{N_R-1} \sum_{x=0}^{N_C-1} s(z, y, x) \cdot \cos(t_1) \cos(t_2) \cos(t_3) \quad (2.1)$$

- 3D-IDCT :

$$s(z, y, x) = \sum_{w=0}^{N_F-1} \sum_{v=0}^{N_R-1} \sum_{u=0}^{N_C-1} a_{3D}(w, v, u) S(w, v, u) \cdot \cos(t_1) \cos(t_2) \cos(t_3) \quad (2.2)$$

여기서

$$t_1 = \frac{(2x+1)u\pi}{2N_C}, \quad t_2 = \frac{(2y+1)v\pi}{2N_R}, \quad t_3 = \frac{(2z+1)w\pi}{2N_F}$$

$$a_{3D}(w, v, u) = \sqrt{\frac{2}{N_F} \frac{2}{N_R} \frac{2}{N_C}} C(w)C(v)C(u)$$

$$\text{그리고, } C(k) = \begin{cases} \frac{1}{\sqrt{2}} & k=0 \\ 1 & k \neq 0 \end{cases}$$

$s(z, y, x)$: 3D 샘플 픽셀 값

(N_F 영상 프레임 ($0 \leq z \leq N_F$))

$S(w, v, u)$: 3D-DCT 계수

($N_F \times N_R \times N_C$ DCT 큐빅블록)

w, v, u : 시간과 수평/수직 주파수

변환영역에 대한 인덱스(Index)

$N_F \times N_R \times N_C$ 구조를 갖는 DCT 큐빅블록은 N_F 영상 프레임들에 대한 각각의 정보를 포함한다. 특히 DCT 내의 큐빅블록에서 첫 번째 프레임 F_0 는 N_F 영상 프레임들에 대한 DC 성분의 정보들을 포함하고 있으며, 두 번째 프레임에서 여덟 번째 프레임($F_1 \sim F_7$)까지의 각 프레임에 대한 대부분의 정보는 저주파에서 고주파 쪽으로 갈수

록 움직임이 많은 성분들을 포함하고 있다.

그러므로 3D-DCT에서 F_0 는 DC 성분으로 8 프레임에 대한 영상 데이터의 각각에 대한 중요한 정보를 포함하고 있음을 보여준다. 나머지 프레임들은 움직이는 고주파 성분의 정보를 포함하고 있으므로 원영상의 시퀀스에서 움직임과 일치한다[11].

영상을 압축할 때 저주파 부분을 고려하게 되면 매우 뛰어난 압축효과를 얻을 수 있다.

영상 블록에 대해 2D-DCT로부터 AC 계수의 분포는 일반적으로 Laplacian 분포(양방향의 지수적 형태) 함수에 의해 모델화 된다. Laplacian 분포 함수는 식 (2.3)과 같이 표현할 수 있다.

$$f(x) = \frac{\lambda}{2} e^{-\lambda|x|} \quad (2.3)$$

식 (2.3)은 zero 값에 대해 대칭적으로 나타난다. 간단한 2D-DCT 블록으로부터 AC 계수의 분포 함수는 파라미터 λ 의 근사 값을 찾음으로서 쉽게 양방향 지수적 분포 함수로 모델링 할 수 있다. 경험적으로 3D-DCT의 AC 계수들에 대한 일반적인 분포에 대하여 이론적 모델로서 등가적 양방향 지수 함수를 사용할 것이다.

그러므로, 이 입체 형태를 보다 적은 영역의 의미있는 계수들을 포함하는 영역에 대하여 점령하는 쌍곡면을 좀 더 정확한 보완성을 갖는 특성을 나타내므로 이것을 이동 보완 쌍곡면(SCH:shifted complement hyperboloid)이라 한다.

III. 실험 결과

그림 3.1의 쌍곡면의 빗금친 영역은 블록을 고려했을 때 쌍곡선에 의해 경계를 이루게 된다.

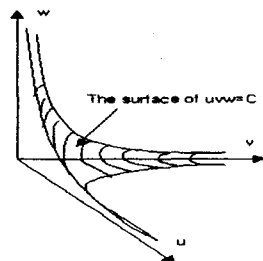


그림 3.1 3차원 그림자면 내의 입체도
Fig. 3.1 A solid diagram under the shaded surface of 3 dimension

사실, AC 계수의 대부분은 전체적 3D-DCT 큐빅블록에 대해 적은 에너지를 포함한다. 이것은 $8 \times 8 \times 8$ 큐빅블록에 대해 511개의 계수로부터 주기적 최고점(peak)을 가지고 있으나 오른쪽으로 갈수록 감소하는 것을 알 수 있다. 이것은 그림 3.2에서 P_1 의 함수로 나타낼 수 있다.

또한, 그림에서 P_2 로 표현되는 함수는 주기가 64인데, 큐빅블록의 각 평면 판에 대해 $u=0, v=0$ 위치에 대응된다. 최소의 주기는 8로서 P_3 에 의한 함수 표현이 가능하고, 각 축에서 3D-DCT 큐빅블록에 대한 시간축을 나타내는 시간차원에 대응된다.

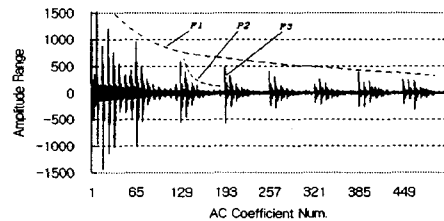


그림 3.2 Pingpong 영상 시퀀스의 AC 계수들의 분포도

Fig. 3.2 AC values distribution diagram of 9-16 frames from Pingpong image sequences

기존의 AC 계수 분포는 그림 3.2에서와 같이 지수 함수적으로 감소하는 형태를 나타내었다.

그러나 본 논문에서 사용된 차분 데이터에 대한 실험 결과, 그림 3.3에서 나타낸 것과 같이 w 값이 커질수록 지수 함수적으로 감소하지 않고 랜덤하게 변한다. 이것은 차분 데이터가 현재영상과 이전영상과의 차이 성분으로 나타나기 때문에 원영상의 데이터보다 적은 엔트로피를 갖고 이웃하는 픽셀간의 상관도 또한 떨어지기 때문이다.

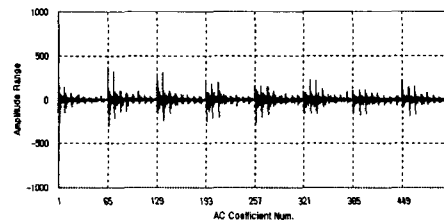


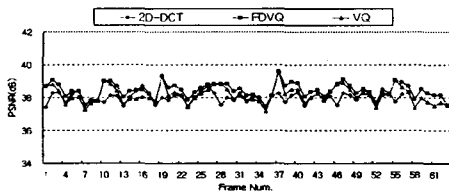
그림 3.3 Pingpong 영상 시퀀스의 1-9 차분 프레임에 대한 AC 계수들의 분포도

Fig. 3.3 AC values distribution diagram of 1-9 diff. frame from Pingpong image sequence

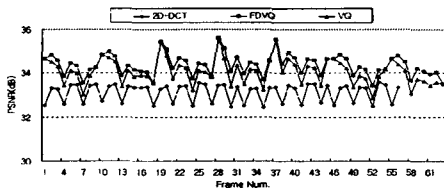
실험결과는 63개의 연속 영상 프레임 시퀀스에 대한 복원 영상들의 평균 PSNR 값을 표 1에 나타내었으며, 각 영상들에 대한 결과 값을 그림 3.4에 나타내었다.

표 1. 시험 동영상들의 PSNR 값
Table 1. The PSNR values of test moving-images

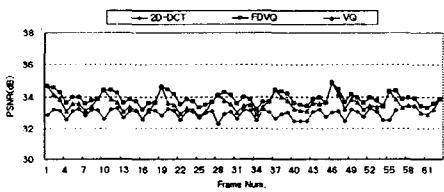
영상	Miss A. (0.048bpp)	Pingpong (0.052bpp)	Child (0.049bpp)	Salesman (0.062bpp)	Lab_img (0.047bpp)	Cp_img (0.050bpp)
FDVQ	38.45dB	30.08dB	34.68dB	33.63dB	34.39dB	33.91dB
VQ	38.16dB	29.65dB	33.99dB	33.39dB	34.08dB	33.55dB
2D-DCT	38.03dB	30.19dB	33.82dB	32.78dB	33.11dB	32.95dB



(a) Miss A.



(b) Lab_img



(c) Cp_img

그림 3.4 63 영상프레임에 대한 복원영상 PSNR
Fig. 3.4 The PSNR of Reconst. image for 63 image frames

IV. 결론

실험 결과, FDVQ 방법에서 기존의 VQ 방법보다 압축률은 거의 같았지만 복원 해상도는 약 0.39dB 향상되었고, 2D-DCT 방법보다 저주파 성

분을 많이 포함하는 동영상 시퀀스(Miss A.)에서는 평균 압축률 0.048bpp에서 0.4dB 정도 향상되었으나, 고주파 성분을 많이 포함하는 동영상 시퀀스(Salesman 등)에서는 평균 압축률 0.052bpp에서 복원 해상도가 0.8dB이상 향상되었다.

결론적으로 고주파 성분을 많이 포함하는 영상 시퀀스에서 뛰어난 복원 영상을 얻을 수 있었으며 높은 압축률에서 실시간 영상 통신을 위한 시스템에 적용할 수 있을 것으로 기대된다.

참고 문헌

- [1] R. J. Clarke, "Transform coding of images", Academic Press, 1985.
- [2] 정제창 역, "최신 MPEG", 교보문고, 1995.
- [3] 호요성, "MPEG 표준화 기술 동향", 한국 정보과학회 논문지, pp.16-26, May, 1996.
- [4] R. DeVore, B. Jarwert, and B. Lucier, "Image compression through wavelet transform coding", *IEEE Transactions on Information Theory*, IT38(2), pp. 719-746, 1992.
- [5] J. W. Woods, "Subband Image Coding", Kluwer Academic, Boston, 1991.
- [6] D. L. Gall, "MPEG: a video compression standard for multimedia applications", *Communication of ACM*, 35, pp.46-58, April, 1991.
- [7] K. Rao and P. Yip, "Discrete Cosine Transform: Algorithm, Advantages, Application", Academic Press, Inc., 1990.
- [8] A. Baskurt and R. Goutte, "3-Dimensional image compression by discrete cosine transform", in *Signal Processing IV: Theories and Applications*, EURASIP, pp.79-82, 1988.
- [9] Farrelle, Paul M., "Recursive Block Coding for Image Data Compression", 1990.
- [10] Joan L. Mitchell, William B. Pennebaker, Chad E. Fogg and Didier J. LeGall, "MPEG Video : Compression standard", Chapman & Hall, 1997.
- [11] Borko Furht, Joshua Greenberg, Raymond West-water, "Motion Estimation Algorithms for Video Compression", Kluwer Academic Publishers, 1997.