

쉬프트와 덧셈을 이용한 DCT 알고리즘

正會員 鄭 華 子* 正會員 金 庠 仲** 正會員 鄭 己 鉉** 正會員 金 容 得**

A DCT Algorithm using shift and additions

Hwa Ja Chung*, Sang Joong Kim**, Gi Hyun Jung**, Yong Deak Kim** *Regular Members*

要 約

영상처리 분야에서 많이 이용되고 있는 離散 코사인 변환(DCT: Discret Cosine Transform) 처리를 위하여 쉬프트와 덧셈으로 곱셈연산을 대체한 새로운 방법을 제안한다. 제안한 방법을 사용하여 DCT 및 IDCT(Inverse DCT)를 거쳐 재생된 영상은 畫質의 劣化를 肉眼으로 感知할 수 없을 정도였으며, PSNR(Peak to peak signal to noise 5ration) 값도 모두 35dB 이상으로서 본 제안의 우수성이 입증되었다.

ABSTRACT

A new approach is proposed for the DCT which is widely utilized in the image processing. The approach replaces mutiplications with shift and additions. In the image restored by the proposed DCT and IDCT, no visible degraation is observed and PSNR(Peak to Peak Signal to Noise Ratio) is greater than 35 dB for all cases, proving the usefulness of the proposal.

I. 序 論

HDTV, G4 FAX, VCR 등의 영상신호를 ISDN(綜合情報通信網: Integrated Service Digital Network)으로 전송하기 위해서는 데이터 압축 기술이 필요하다. DCT는 最適變換으로 알려진 KLT(Karhunen Loeve Transform)의 성능에 가장 근접한 성

능을 가지고 있으며, 정보가 한 곳으로 집중되는 효과가 크기 때문에 영상정보와 음성정보의 부호화에 많이 사용된다. 특히 JPEG(Joint Photographic Expert Group), MPEG(Motion Picture Expert Group), H.261(TV 會議/電話符號化標準) 등 國際標準規格에서 DCT를 기본 알고리즘으로 채택되고 있다.

DCT는 벡터의 각 성분(orthogonal base)을 코사인 함수로 표현하며, N차(행렬 크기 $N \times N$) 변환 행렬 $[A]$ 에서 i행 j 열의 요소 a_{ij} 는 다음과 같은 식으로 주어진다.⁽¹⁾

*서울産業大學校 電子計算學科
**亞州大學校 電子工學科
Dept. of Electronic Engineering Ajou University.
論文番號: 93-77

i = 0일 때,

$$a_{ij} = \frac{1}{\sqrt{N}} \quad (\text{식 1.a})$$

i ≠ 0일 때,

$$a_{ij} = \sqrt{\frac{2}{N}} \cos \frac{(2j+1)i\pi}{2N} \quad (\text{식 1.b})$$

한편, n차원 평면에서의 좌표 변환은

$$[Y] = [A][X] \quad (\text{식 2})$$

로 표시되므로, 식(1.a) (1.b)를 식(2)에 대입하여 식(2)의 행렬을 전개하면, 1차원 N차 변환 및 그 逆變換은 식(3)과 식(4)로 표현된다.

$$T_i = \sqrt{\frac{2}{N}} \alpha_i \sum_{j=0}^{N-1} X_j \cos \frac{(2j+1)i\pi}{2N} \quad ; i = 0, 1, \dots, N-1 \quad (\text{식 3})$$

$$X_i = \sqrt{\frac{2}{N}} \sum_{j=0}^{N-1} \alpha_j T_j \cos \frac{(2j+1)i\pi}{2N} \quad ; i = 0, 1, \dots, N-1 \quad (\text{식 4})$$

$$\text{단, } \alpha_0 = \frac{1}{\sqrt{2}} \\ \alpha_i = 1 \quad ; i \neq 0$$

한편, 2차원 N × N FDCT 및 IDCT는 식(5) 및 식(6)으로 표현될 수 있다.

$$T_{uv} = \frac{2}{N} C_u C_v \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} G_{xy} \cos \frac{(2x+1)u\pi}{2N} \cos \frac{(2y+1)v\pi}{2N} \quad (\text{식 5})$$

$$G_{uv} = \frac{2}{N} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} C_u C_v T_{uv} \cos \frac{(2x+1)u\pi}{2N} \cos \frac{(2y+1)v\pi}{2N} \quad (\text{식 6})$$

$$\text{단, } C_u, C_v = \frac{1}{\sqrt{2}} \quad ; u, v = 0$$

$$C_u, C_v = 1 \quad ; u, v \neq 0$$

FDCT 및 IDCT는 시간이 많이 소요되는 연산이면서도 압축 및 복원시 여러번 실행되는 연산이다. 예를 들어, 크기 640 × 400 화소의 영상을 8 × 8 블록으로 된 화소로 분해하여 처리하는 경우 640 × 400 ÷ 64 = 4000 회의 FDCT 및 IDCT가 실행된다. 그러므로, FDCT 및 IDCT의 처리 속도를 높이는 것은 영상 압축/복원 과정의 소요시간을 줄이는 것과 직결되는 일이다.

지금까지의 고속 DCT 연산 방법들은 FFT(Fast Fourier Transform)나 다른 변환 방법을 이용하는 간접 방법⁽²⁾과 DCT의 변환 행렬을 분해하는 직접적인 방법⁽²⁾⁽³⁾으로 대별할 수 있다. 1974년 Ahmed⁽²⁾ 이래 행렬을 분해하여 곱셈 횟수를 줄이는 고속 알고리즘들⁽³⁾⁻⁽⁶⁾이 연구되어 왔다. 행렬 분해 방법들의 주된 목적은 DCT 계산 과정에서의 곱셈횟수를 줄이는 것이다. 왜냐하면, 곱셈회수의 감소는 처리속도의 향상과 직결되기 때문이다. 그러나, 이들 고속 알고리즘에서는 어느 경우이든 최소한의 곱셈 연산을 필요로 하는 공통점이 있다. 더우기 N × N 2차원 영상의 DCT 처리를 위해서는 변환 행렬이 N² × N² 이 되므로 곱셈횟수는 기하급수적으로 증가하여 급격한 처리속도의 저하를 가져온다. 2차원 변환 행렬의 분해에 관한 것으로는 Haque⁽⁵⁾, Hou⁽⁶⁾ 등이 제안한 방법들이 있다.

본 논문에서는 이러한 고속 알고리즘을 실현하기 위하여 지금까지의 방법과는 달리 FDCT 및 IDCT의 연산과정에서 처리 속도에 직접적으로 영향을 미치는 곱셈연산을 쉬프트 및 덧셈연산으로 대체함으로써 곱셈연산으로 인한 문제를 근본적으로 해결할 수 있는 방법을 제안한다. 본 제안 방법을 적용하여 표준 영상을 처리하고 그 결과를 코사인 값을 직접 곱하여 처리한 결과와 비교하였다. 본 제안 방법으로 처리된 결과는 畫質의 劣畫를 可視的으로 感知하기 어려웠으며 PSNR 값도 코사인 값을 직접 곱했을 때의 결과에 近接觸을 볼 수 있다. 이는 제안 알고리즘에 내포하고 있는 非可逆性이 화질에는 크게 영향을 미치지 않음을 실험을 통하여 입증하는 것이다.

II. 제안 알고리즘

8 × 8의 2차원 영상의 DCT를 위한 변환 행렬에 나타나는 코사인 값의 종류는 부호를 고려하지 않을 경우 cos 1π/16, cos 2π/16, ... cos 7π/16까지의 7가지이다. 이들중 cos 5π/16를 선택하여 기준으로 삼고,

나머지 값들을 그 기준값에 대해 정규화한다. 이 정규화된 값들은 2진 코드로 표시된다. 한편, 입력 화소값은 수 차례에 걸쳐 오른쪽으로 1비트씩 쉬프트되며, 각각의 쉬프트된 값들은 해당 레지스터들에 저장된다. 이들 레지스터 내에 저장된 쉬프트된 입력 값들은 위의 정규화된 2진 코드에 따라 선택적으로 더해진다. 이 경우, 정확도는 쉬프트 횟수의 증가에 따라 향상된다. 그러나, 쉬프트 횟수의 증가는 처리 속도의 저하를 초래하므로 쉬프트 횟수의 적절한 선택이 필요하다. 단, 이상의 처리를 하기에 앞서, 기준으로 삼은 코사인 값을 1로 정규화한 값에 해당하는 이진 코드 값을 입력 화소값에 대해 위와 같은 방법으로 곱셈을 대체해 주어야 한다.

1차원 $N \times 1$ 차 DCT를 數式으로 표현하면 아래와 같다. 편의상 $\sqrt{2}/N$ 및 $2C_u C_v/N$ 의 상수 부분(scale factor)을 생략하고 식(7)에서 식(10)까지와 같이 정의되는 定規化(normalization)된 DCT로 나타내었다.

$$T_m = \sum_{v=0}^{N-1} F_v \cos \frac{(2v+1)m\pi}{2N} \quad (\text{식 7})$$

$$G_m = \sum_{v=0}^{N-1} \left(\sum_{i=0}^{M-1} a_i \cdot RS(F_v)_i \right) \quad ; M : \text{쉬프트 횟수} \quad (\text{식 8})$$

단 a_i 는 i 번째 비트 코드이며, $RS(F_v)_i$ 는 F_v 을 오른쪽으로 i 비트 쉬프트시키는 함수이다.

이를 2차원 $N \times N$ DCT에 대해 나타내면 아래와 같이 표현될 수 있다.

$$T_m = \sum_{v=0}^{N-1} F_v \cos \frac{(2v+1)u\pi}{2N} \quad (\text{식 9})$$

$$G_m = \sum_{v=0}^{N-1} \left(\sum_{i=0}^{M-1} a_i \cdot RS(F_v)_i \right) \quad (\text{식 10})$$

단, b_j 는 j 번째 비트 코드이다.

$N=8$, $M=8$ 이고, $\cos 5\pi/16$ 를 기준으로 했을 때, 입력 화소값 F_{xy} 에 $\cos 1\pi/16$ 를 곱한 경우를 예로 들어 설명하면 다음과 같다.

우선, $\cos 1\pi/16$, $\cos 2\pi/16$, ..., $\cos 7\pi/16$ 의 값을 각각 $\cos 5\pi/16$ 으로 나누고, 이를 2진 코드로 표시하여 표1을 얻는다.

표1. $N=8$, $M=8$ 일 때 $\cos 5\pi/16$ 로 정규화된 코사인 값 및 2진 표시

Table 1. Normalized cosine values and binary codes with $N=8$, $M=8$ and $\cos 5\pi/16$ for the base

n	원 래 값	$\cos 5\pi/16$ 기준	2진 표시
1	0.9807852804	1.7653668647	1.1100001
2	0.9238795325	1.6629392246	1.1010100
3	0.8314696123	1.4966057627	1.0111111
4	0.7071067912	1.2727585806	1.0100010
5	0.5555702330	1.0000000000	1.0000000
6	0.3826834324	0.6888119802	0.1011000
7	0.1950903222	0.3511533024	0.0101100

먼저, 입력 화소값 F_{xy} 를 $\cos 5\pi/16$ 기준으로 변환하기 위하여 0.5555702330배로 level down한 후 레지스터 R0에 넣는다. 이 과정 역시 $SR(F_{xy})_1 + SR(F_{xy})_5 + SR(F_{xy})_6 + SR(F_{xy})_7 + SR(F_{xy})_{11} + SR(F_{xy})_{12} + SR(F_{xy})_{13}$ (0.5555702330는 0.1000111000111B로 coding된다.)의 쉬프트 및 덧셈연산을 이용한다. 다음에, R0를 오른쪽으로 1비트씩 계속 쉬프트해가며 각각의 값을 레지스터 R1, R2, ..., R7에 차례로 넣은 후, 그 2진 표시 중에서 비트 값이 '1'인 비트 번호에 대응되는 레지스터들의 값을 더한다. 즉, $F'_{xy} \times \cos 1\pi/16$ 는 $R0 + R1 + R2 + R7$ 이 된다. 이렇게 하여 구해진 값은 원래의 값에 비해 약 0.4%의 오차를 가진다(다른 코사인 값에서는 최대 2.1%, 최소 0.2%의 오차를 가진다). 제안된 방법으로 한 계산에서 발생하는 이러한 구조적인 오차는 DCT 과정을 非可逆化시킨다. 그러나, 다음 절의 실험을 통해서 이러한 오차가 실제의 영상 처리 과정에서는 화질에 크게 영향을 미치지 않음을 알 수 있어 제안 알고리즘의 유용성이 입증된다.

III. 실험 및 결과

FDCT 및 IDCT를 II에서 제안한 방법을 적용하여 실험해 보았다. 본 실험에서는 입력 영상은 256 grey level을 갖는 256×256 화소의 표준 영상 LENA를 사용하였으며, JPEG⁽⁷⁾에서와 같이 8×8 화소로 된 블록 단위의 grey scale 영상정보를 처리의 기본 단위로 정하여 FDCT 및 IDCT를 한 후, 원화상으로 복원하였다. 식(11)과 식(12)는 본 실험에서 사용한 8×8 FDCT와 8×8 IDCT의 수학적 정의를 나타낸다.

$$T_{uv} = \frac{1}{4} C_u C_v \sum_{x=0}^7 \sum_{y=0}^7 (G_{xy} - L_s) \cos \frac{(2x+1)u\pi}{16} \cos \frac{(2y+1)v\pi}{16} \quad (\text{식 10})$$

$$G_{xy} = \frac{1}{4} \sum_{u=0}^7 \sum_{v=0}^7 C_u C_v (T_{uv} + L_s) \cos \frac{(2u+1)x\pi}{16} \cos \frac{(2v+1)y\pi}{16} \quad (\text{식 11})$$

단, $C_u, C_v = \frac{1}{\sqrt{2}}$; $u, v = 0$
 $= 1$; 기타
 $L_s = 128$; P_{xy} 의 비트 정밀도 = 8 비트

식(11)에서 각 화소값에서 L_s 값을 빼는 것은 DC 계수의 기대값을 0으로 레벨쉬프트하기 위한 것이고, 식(12)에서 L_s 값을 더해 주는 것은 원래 값으로 복원하기 위한 것이다. 식(11) 및, 식(12)에서 C_u 및 C_v 값은 u 및 v 값이 0이 아닌 경우에는 1이 되므로 무시해도 되지만, 0인 경우는 $\frac{1}{\sqrt{2}}$ 이 되므로 이것을 그 값에 해당되는 $\cos 4\pi/16$ 로 간주하여 처리하였다.

제한 알고리즘의 효율성 및 정확도는 쉬프트 레지스터의 수(이하 M이라 한다) 및 입력 화소값을 왼쪽으로 레벨 쉬프트 해 주는 비트 수(이하 SLL이라 한다)에 영향을 받는다. SLL은 쉬프트 동작 도중에 발생할 수 있는 오른쪽 절사를 방지하기 위함이다 그 목적이며, 가능한 한 절사를 막음으로써 정확도를 높일 수 있다. 그림 1은 여러가지 값의 M과 SLL에 따른 실험 결과를 나타낸 것이다.

$M < 4$ 인 경우 오차가 너무 크며, $M > 10$ 인 경우 PSNR 값이 크게 변하지 않았다. SLL < 11에서는 PSNR 값이 급격히 감소하고, SLL 값이 21보다 큰 범위에서는 실험 장비의 제약으로 인해 실험이 불가능하였다(그러나, 이는 제한 알고리즘의 성능측정에 크게 영향을 미치지 않음을 실험 결과를 통해 알 수 있다).

실험결과에 대한 평가는 객관적 화질 평가 방법의 하나인 식(13)의 PSNR을 사용하였다.

$$PSNR = 10 \log_{10} \{ [\text{입력 화소값 } F(x, y) \text{ 중 퍼크값}]^2 \div E\} \text{ dB}$$

$$E = \frac{1}{255^2} \sum_{x=0}^{255} \sum_{y=0}^{255} (G(x, y) - F(x, y))^2 \quad (\text{식 13})$$

그림 1에 나타난 바와 같이 실험한 모든 경우에 PSNR 값이 35dB 이상으로 눈에 띄는劣質을 식별하기 힘들었으며 블러화 현상도 볼 수 없었다. 특히 $M > 7$ 인 경우는 모두 PSNR이 42 dB를 넘어 코사인 값을 직접 곱했을 경우의 약 48 dB에 근접함을 알 수 있다(또 다른 표준 영상의 하나인 PEPP의 경우에서도 유사한 PSNR 값을 얻었다). 그림2와 그림3은 원 화상과 제안 알고리즘으로 압축/복원한 영상의 예이다. 예상했던 바와 같이 SLL의 증가는 PSNR의 증가 요인이 됨을 알 수 있으나 전체적인 PSNR에는 크게 영향이 없음을 알 수 있다. 이는 SLL이 작아서 절사가 발생하게 되는 부분은 계산치의 ISB에 해당되는 것으로 전체값에서 적은 비중을 차지하기 때문이며, 절사가 발생하지 않을 정도로 SLL이 큰 경우에는 SLL이 증가해도 정확도는 증가하지 않음에 기인한 것이다. M의 값은 증가함에 따라 PSNR도 증가한다. 이는 M의 증가가 제한 알고리즘에서 구현되는 擬似곱셈값을 原 곱셈값에 접근시키기 때문이다. 그림1에서 M의 값 10 이상에서 PSNR이 감소하기 시작하는 것은 본 실험에 사용한 장비의 制約으로 인한 것이다. M이 10을 초과하는 경우 오른쪽 절사를 유발하여 M의 증가가 오히려 정확도의 감소를 초래하기 때문이다. 그러나 이러한 절사에도 불구하고 PSNR 값은 그다지 감소하지 않음을 알 수 있다.

제한한 방법과 코사인 값을 직접 곱한 방법으로 영

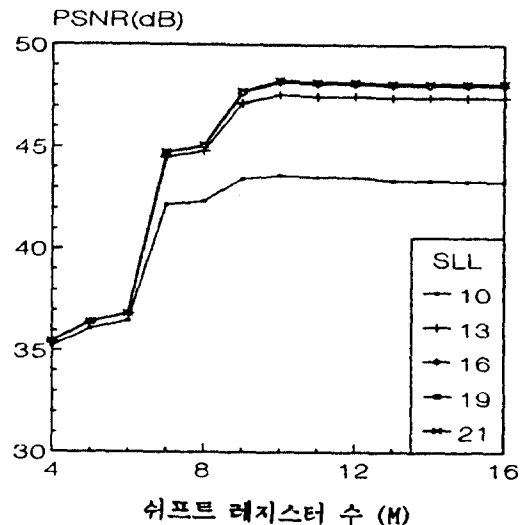


그림 1. M 및 SLL에 따른 PSNR
 Fig 1. PSNR for different values of M and SLL.



그림 2. 원 영상 Lenna
Fig 2. Original Image of Lenna



그림 3. 제안 방식에 의해 재생된 영상
Fig 3. Reconstructed Image by the proposed approach



그림 4. 코사인 값을 직접 곱한 방식에 의해 재생된 영상
Fig 4. Reconstructed Image by the cosine multiplication

상을 처리하는데 소요되는 실행시간의 비는 i80386 PC에서 약 1:7.3이었다(coprocessor를 사용하여 코사인 곱셈의 속도를 향상 시킨 경우도 약 1:1.2). 이는 제안 알고리즘이 처리속도의 향상에 유용함을 보여 주는 것이다.

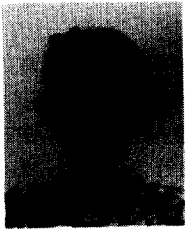
IV. 結 論

본 논문에서는 영상처리 분야에서 많이 이용되고 있는 DCT 처리를 위한 새로운 방법을 제안하였다. 많은 처리 시간을 요하는 곱셈 연산을 덧셈 및 쉬프트 연산으로 처리하기 위하여 非可逆의 近似法이 제시되었다. 제안된 방법의 효율성을 표준영상의 실험을 통하여 입증하였으며 코사인 값을 직접 곱하는 방식에 의한 처리 결과와 비교되었다. 再生된 영상은 화질의 열화를 肉眠으로 감지하기 어려웠으며 PSNR 값도 대부분의 경우 47 dB 이상을 나타냄으로서 기존의 방법으로 얻어진 48.51 dB와 큰 차이를 보이지 않아 제안된 방법의 유용성이 입증되었다. 본 제안 방식을 알려진 기존의 여러 고속 알고리즘들에 적용할 경우 적용 알고리즘의 성능에 비례하여 더 큰 성능 향상을 기대할 수 있다.

參 考 文 獻

1. R.Rao, P.Yip, Discrete Cosine Transform Algorithms, Advantages, Applications New York, Academic press, 1990.
2. M.Vetterli, H.Nussbaumer, "Simple FFT and DCT algorithms with reduced number of operations," Signal Processing, vol.6, pp.267-278, Aug. 1984.
3. N.Ahmed, T.Natarajan, K.Rao, "Discrete cosine transform," IEEE Trans. Comput., vol.C-23, pp.90-93, Jan. 1984.
4. B.G.Lee, "A new algorithm for the discrete cosine transform," IEEE Trans, Acoust., Speech, and Signal Process., vol. ASSP-32, No.6, pp. 1243-1245, Dec.1984.
5. M.A.Haque, "A two-dimensional fast cosine transform," IEEE Trans. Acoust., Speech, and Signal Process., vol. ASSP-33, No.6, pp.1532-1538, Dec. 1985.

6. H.S.Hou, "A fast recursive algorithm for computing the cosine transform," IEEE Trans Acoust., Speech, and Signal Process., vol. ASSP-35, No.10, pp.1455-1461, Oct 1987.



鄭 華 子(Hwa Ja Chung) 正會員
1947年 8月 3日生
1971년 : 연세대학교 문학사(문헌정보학)
1983년 : 연세대학교 공학석사(전자계산학)
1993년 : 아주대학교 공학박사(전자공학)

1985년 ~ 현재 : 서울산업대학교 전자계산학과 부교수
※주관심분야 : HCI, 멀티미디어, 영상압축



金 庠 仲(Sang Joong Kim) 正會員
1949년 5월 15일생
1977년 2월 : 한양대학교 전자공학과 학사
1980년 8월 : 연세대학교 대학원 전자공학과 석사
1992년 3월 ~ 현재 : 아주대학교 대학원 전자공학과 박사과정

1977년 12월 ~ 현재 : 한국전자통신연구소 통신접속연구실장
※주관심분야 : 통신시스템 및 영상처리 등



鄭 己 鉉(Gi Hyun Jung) 정회원
1958년 10월 21日生
1984년 : 서강대학교 전자공학과 졸업
1988년 : Univ. of Illinois 공학석사
1990년 : Purdue Univ. 공학박사
1984년 ~ 1986년 : FujiTsu Korea System Engineer

1991년 ~ 1992년 : 현대전자 반도체연구소 책임연구원
1992년 ~ 현재 : 아주대학교 전자공학과 조교수
※주관심분야 : 컴퓨터구조, VLSI 설계, ASLC, Multi media 등

金 容 得(Yong Deak Kim)

正會員

1946년 1월 30日生

1971년 : 연세대학교 전자공학과 졸업
1973년 : 연세대학교 대학원(공학석사)
1978년 : 연세대학교 대학원(공학박사)
1973년 ~ 1974년 : 불란서 ESE 연구원
1979년 ~ 1980년 : 미국 Stanford대학교 연구교수 재직
1978년 ~ 현재 : 아주대학교 전자공학과 교수

※주관심분야 : 디지털 시스템에 관련된 하드웨어 응용, 특히 FA, OA 네트워크 응용과 접속방안, 멀티미디어 시스템 구축과 화상 전송방식, ISDN망 접속 등에 흥미를 갖고 있음.