

DCT 영역에서의 분류화 방법을 이용한 벡터 양자화기

°임창훈, 고종석, 김계균
한국과학기술원 전기 및 전자

A new Classified VQ Algorithm in the DCT domain

Chang-Hoon Yim, Jong-Seog Koh, Jae-kyoon Kim
Dept. of Electrical Engineering, KAIST

ABSTRACT

In this paper, we propose a new Classified Vector Quantization algorithm in the DCT domain (DCT-CVQ). DCT sub-blocks are classified simply into 6 classes using two DCT coefficients which well represent edge characteristics, and VQ codebook size for each class is determined by its ac energy level. It is shown that, at low bit rates, the subjective quality is satisfactable since both the edge degradation and the mismatch effect are significantly reduced.

1. 서론

VQ(Vector Quantization)는 낮은 bit rate에 적합한 coding 기법으로 여러가지의 장점을 가지고 있는 반면 몇가지의 단점을 가지고 있다. 그것은 distortion measure로 MSE(mean square error)가 쓰일 때 나타나는 edge degradation과 bit rate와 block size가 증가함에 따라 지수적으로 증가하는 계산량, 그리고 codebook을 설계할 때 training vector로 사용한 image와 coding되는 image가 다를 때 성능이 떨어지는 mismatch 현상 등이다.

이와같은 VQ의 문제점을 해결하기 위한 여러가지 VQ 방식들이 제안되었는데[2], 최근에는 edge degradation과 계산량 문제에 효과적인 CVQ(Classified Vector Quantization)가 제안되었다.[3] 이것은 edge의 방향성에 따라서 class를 나누고, 각 class마다 따로 codebook을 design 하는 방식으로, 이런 CVQ는 edge degradation이 더 적고, 각 class의 codebook size는 전체의 codebook size보다 훨씬 작으므로 encoding 할 때의 계산량이 훨씬 작다.

최근에는 이와 같은 spatial domain에서의 CVQ와 마찬가지로 transform domain에서의 CVQ도 연구가 진행되고 있는데, DCT 변환계수의 energy 집중 부분에 의해서 classification 하는 방식과[4], block을 edge 특성에 의해서 4개의 category로 나누어서 VQ하는 방식[5] 등이 제안되었다.

본 논문에서는 edge의 특성을 이용한 DCT domain에서의 classification 방식으로, edge와 DCT 변환계수들과의 관계에 의해서 class를 나누고, Chen 등이 제안한 DCT domain에서의 adaptive coding 기법[6]을 도입한 DCT-CVQ (DCT Classified Vector Quantization) 방식을 제안하고자 한다.

2. Edge와 DCT 변환계수와의 관계

본 절에서는 DCT domain에서 edge의 특성을 고려한 효과적인 classification을 하기 위해 먼저 edge와 DCT 변환계수와의 관계를 살펴본다.

2D DCT를 matrix에 관한 식으로 나타내기 위해 변환계수 matrix를 $[a]$, 원래의 image block matrix를 $[G]$, 1D DCT transform matrix를 $[T]$ 로 나타내면, 2D DCT 식은 다음과 같다.

$$[a] = [T][G][T]^T \tag{1}$$

4×4 DCT의 경우 $[T]$ 는 다음과 같다.[8]

$$[T] = \frac{1}{2} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \cos\frac{\pi}{8} & \cos\frac{3\pi}{8} & -\cos\frac{3\pi}{8} & -\cos\frac{\pi}{8} \\ \cos\frac{2\pi}{8} & -\cos\frac{2\pi}{8} & -\cos\frac{2\pi}{8} & \cos\frac{2\pi}{8} \\ \cos\frac{3\pi}{8} & -\cos\frac{\pi}{8} & \cos\frac{\pi}{8} & -\cos\frac{3\pi}{8} \end{bmatrix} \tag{2}$$

(1)에서 정의한 DCT 변환계수 중 α_{00} 는 원래의 image block 내의 DC값을 나타내고, α_{01} 과 α_{10} 는 edge의 특성을 가장 잘 나타낸다.

먼저 vertical edge와 α_{01}, α_{10} 의 관계를 알아보기 위해 4x4 image block 내에 vertical edge가 있는 경우에, 그림 1의 (a),(b),(c)의 3가지 경우가 있다고 가정한다.

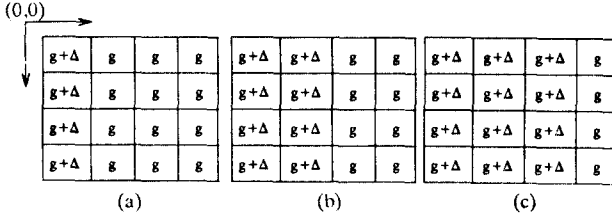


그림 1. Vertical edge가 있는 경우의 image block (4x4)

그림 1에서 g는 gray level을 나타내고, Δ는 edge에서의 gray level의 변화량을 나타낸다. 그리고 Δ의 부호는 edge의 극성, 즉 block 내에서의 gray level의 변화방향을 나타내는데, 그림 1에서는 왼쪽에서 오른쪽으로 감에 따라 gray level이 Δ>0이면 작아지고, Δ<0이면 커지는 것을 나타낸다. 그림 1의 (a),(b),(c) 경우에 각각 α_{01}, α_{10} 를 계산하면,

$$(a) \alpha_{01} = \frac{1}{4\sqrt{2}} \Delta 4 \cos \frac{\pi}{8}, \alpha_{10} = 0$$

$$(b) \alpha_{01} = \frac{1}{4\sqrt{2}} \Delta (4 \cos \frac{\pi}{8} + 4 \cos \frac{3\pi}{8}), \alpha_{10} = 0$$

$$(c) \alpha_{01} = \frac{1}{4\sqrt{2}} \Delta 4 \cos \frac{\pi}{8}, \alpha_{10} = 0$$

이고 α_{01}, α_{10} 의 기대값과 제곱의 기대값을 구해보면 각각 다음과 같다.

$$E[\alpha_{01}] = 0.743\Delta, E[\alpha_{10}] = 0 \quad (3a)$$

$$E[\alpha_{01}^2] = 0.569\Delta^2, E[\alpha_{10}^2] = 0 \quad (3b)$$

따라서 vertical edge가 있는 경우에는 $\alpha_{10} = 0$ 이고, α_{01} 은 Δ에 비례하는데, Δ>0이면 $\alpha_{01} > 0$ 이고, Δ<0이면 $\alpha_{01} < 0$ 인 것을 알 수 있다. 같은 방법으로 horizontal edge의 경우에도 α_{01} 과 α_{10} 와의 관계를 알 수 있다. 위의 결과를 그림으로 나타내면 그림 2와 같은데, 여기에서 빗금친 부분은 gray level이 어두운 부분을, 빗금친 없는 부분은 밝은 부분을 나타낸다.

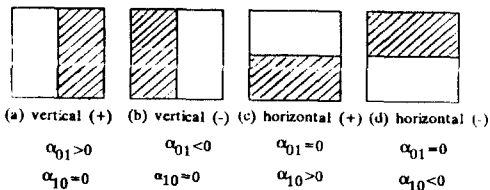


그림 2. Image block 내에 vertical, horizontal edge와 DCT 변환계수 α_{01}, α_{10} 의 관계

마찬가지로 diagonal edge와 α_{01}, α_{10} 와의 관계도 위에서와 같은 방법으로 알아낼 수 있는데, 그 결과를 나타내면 그림 3과 같다.

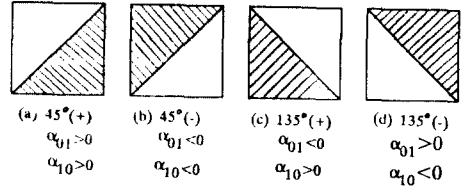


그림 3. Image block 내의 diagonal edge와 DCT 변환계수 α_{01}, α_{10} 의 관계

이상에서 α_{01}, α_{10} 의 두 변환계수에 의하여 edge의 방향과 극성, 그리고 edge에서의 gray level의 변화량을 알 수 있음을 살펴보았다.

다음으로 edge와 나머지 변환계수들과의 관계를 알아보기 위해, 위와같은 방법으로 DC값을 나타내는 α_{00} 을 제외한 나머지 변환계수의 기대값을 구해보면 항상 0인 것을 알 수 있고, 모든 edge의 경우에 α_{00} 을 제외한 나머지 변환계수의 분산을 Δ^2 으로 나눈값을 구해보면 표1과 같은 관계를 얻는다. 여기에서 α_{01}, α_{10} 을 제외한 나머지 변환계수는 diagonal edge에서 edge의 각도에 관계없이 그 변환계수의 특성이 같고, 모든 edge의 경우에 있어서 변환계수의 특성이 edge의 극성에는 관계가 없는 것을 알 수 있다.

표1에서 분산은 변환계수의 ac energy를 나타내므로, 표1에서 각 변환계수의 분산을 보면, vertical edge의 경우에는 윗부분의 변환계수에 energy가 집중적으로 모이고, horizontal edge의 경우에는 왼쪽의 변환계수에 energy가 집중적으로 모이는 것을 알 수 있다. 이에 비해 diagonal edge의 경우에는 좌측상단의 변환계수에 energy가 모이지만, ac energy가 전체적으로 커지는 것을 알 수 있다.

이상에서 4x4 block의 경우의 DCT 변환계수와 edge와의 관계를 살펴보았는데, 8x8 block의 경우에도 위에서 설명한 4x4 block의 경우와 거의 비슷한 특징을 갖는다.

※ 1. 각 edge 경우에 대한 DCT 변환계수들의 분산/Δ² (4x4)

(0,0)	(0,1)	(0,2)	(0,3)	(1,0)	(1,1)	(1,2)	(1,3)	(2,0)	(2,1)	(2,2)	(2,3)	(3,0)	(3,1)	(3,2)	(3,3)
•	.57	.17	.10	•	0	0	0	•	.16	.01	.002	•	.16	.01	.002
0	0	0	0	.57	0	0	0	0	.16	.09	.03	.005	0	0	0
0	0	0	0	.17	0	0	0	.01	.03	.03	.01	0	0	0	0
0	0	0	0	.10	0	0	0	.002	.005	.01	.03	0	0	0	0

(a) vertical

(b) horizontal

(c) diagonal

3. Edge 특성에 의한 DCT domain에서의 classification

2절에서 살펴본 바와 같이, DCT 변환계수와 edge와의 관계를 이용하여 변환계수 α_{01}, α_{10} 에 의하여 4개의 class로 나눌 수 있다.

(1)Homogeneous (2)Vertical (3)Horizontal (4)Diagonal

그런데 diagonal edge는 45° 와 135° 의 각도를 갖는 2가지 경우를 포함하고, 표 1에서와 같이 변환계수의 energy가 전체적으로 커지기 때문에 class를 더 세분할 필요가 있다. Diagonal edge는 변환계수의 특성이 edge 각도와 구성에는 무관하기 때문에 ac energy에 따라 나누면 효과적인데, α_{01}, α_{10} 의 절대값에 의하여 나누면, 전체 ac energy에 의하여 class를 나누는 것과 거의 비슷한 효과가 있다. Diagonal edge를 몇 개의 class로 세분하느냐 하는 것은 bit rate와 image의 성능을 고려하여 실험적으로 결정하였는데, 여기에서는 2개의 class로 세분하였다. 그리고 homogeneous class는 edge 성분이 거의 없는 shade와 약간의 edge 성분이 있는 midrange class로 세분하였다. 이와같이 shade, midrange, vertical edge, horizontal edge와 2개의 Diagonal edge의 6개의 class로 나누는 경우의 classification 과정을 나타내면 그림 4와 같다.

$d = \max(|\alpha_{01}|, |\alpha_{10}|)$
 t_1 : shade와 midrange를 구분하는 threshold
 t_2 : homogeneous와 edge를 구분하는 threshold
 t_3 : diagonal edge를 세분하는 threshold
 λ : horizontal, vertical edge와 diagonal edge를 구분하는 threshold

```

IF  $d < t_1$  THEN
  class = 1 (shade)
ELSE IF  $d < t_2$  THEN
  class = 2 (midrange)
ELSE IF  $|\alpha_{01}| > t_2$  and  $|\alpha_{10}/\alpha_{01}| < \lambda$  THEN
  class = 3 (vertical)
ELSE IF  $|\alpha_{10}| > t_2$  and  $|\alpha_{01}/\alpha_{10}| < \lambda$  THEN
  class = 4 (horizontal)
ELSE IF  $d < t_3$  THEN
  class = 5 (diagonal 1)
ELSE
  class = 6 (diagonal 2)
END IF
  
```

그림 4. DCT domain에서의 classification

4. DCT-CVQ의 구성

DCT-CVQ를 할 때, low bit rate로 하기 위해서는 block size가 4×4 보다는 8×8 경우가 더 적절한데, 그 이유는 4×4 경우의 DCT는 8×8 경우보다 energy 감축 효과가 작고, DC 값에 대한 bit 할당과 class 전송에 필요한 bit가 8×8 경우보다 상대적으로 크기 때문에 낮은 bit rate에서는 사용하기 곤란하다. 따라서 여기에서는 8×8 block에 대해서 DCT-CVQ를 했는데 8×8 block에서 zonal sampling에 의하여 vector를 만드는 방법은 그림 5와 같다.[7]

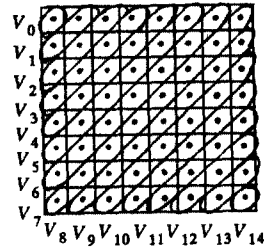


그림 5. 변환계수들의 벡터화 (8×8)

그림 5에서 14개의 vector 중 V_0 는 DC 값이므로 uniform scalar quantizer를 사용하여 coding 하고, energy가 작은 vector는 truncation 한다. VQ해서 전송하는 vector를 V_1, V_2, \dots, V_L 이라고하고, class 수를 M이라 할 때, DCT-CVQ의 송신단 구성을 나타내면 그림 6과 같다. 여기에서 C_{ji} 는 class i, vector j에 대한 codebook을 나타낸다.

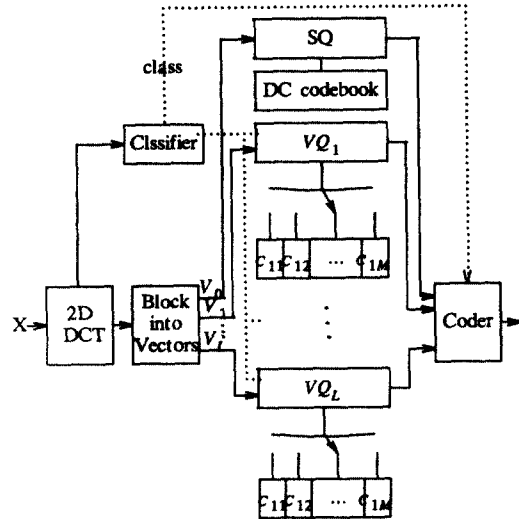


그림 6. DCT-CVQ의 송신단 구성도

DCT block이 class i일 확률을 p_i , class i에서 DC coefficient를 제외한 모든 vector의 index를 전송하는데 필요한 bit를 B_i , DC 값을 전송하는데 필요한 bit를 B_{DC} , class를 전송하는데 필요한 bit를 B_C 라 하면, 한 block에 필요한 평균 bit B_{av} 는 다음과 같다.

$$B_{av} = \sum_{i=1}^M p_i B_i + B_{DC} + B_C \quad (4)$$

식 (4)에서 B_C 는 효과적인 encoding 방법인 Huffman coding[9] 방법을 이용하면 bit를 줄일수 있는데, 6 class인 경우 2.5 bit 미만으로 줄일 수 있다.

각 class에 대해서 codebook을 설계할 때에는 각 vector에 대해서 할당하는 bit수를 ac energy에 의하여 각각 다르게 조절한다. 즉, ac energy가 큰 class에 대해서는

codebook size를 크게하여 정확한 재생을 하고, ac energy가 작은 class에 대해서는 codebook size를 작게하여 효과적인 coding이 되도록 한다.

Class i 에 대해서 (u,v) 번째 DCT 변환계수의 분산을 $\sigma_i^2(u,v)$ 라 하면, 거기에 할당되는 bit 수는 rate distortion theory에 의하여 다음과 같이 구한다.[6]

$$N_{B_i}(u,v) = \frac{1}{2} \log_2 [\sigma_i^2(u,v) - \log_2 D] \quad (5)$$

$$(u,v) \neq 0, u,v = 0,1,\dots,7$$

그리고 class i , vector \underline{V}_i 에 할당되는 bit 수를 b_{ji} 라 하면 여기에 할당되는 codebook size는 $2^{b_{ji}}$ 가 되고,

$$b_{ji} = \sum_{(u,v) \in \underline{V}_i} N_{B_i}(u,v) \quad (6)$$

$$B_i = \sum_{j=1}^L b_{ji} \quad (7)$$

이 된다.

여기에서 모든 ij 에 대해서 할당되는 bit 수 b_{ji} 를 결정하기 위해서는 식(5)에서 D 를 적절히 조절하여, (4),(5),(6),(7) 식에서 구한 B_{av} 가 원하는 bit가 되도록 한다. 그런데 여기에 해당하는 codebook size는 $2^{b_{ji}}$ 이므로, 실제적으로 memory와 계산량 때문에 b_{ji} 를 10 이상으로 하기는 곤란하여, b_{ji} 가 10 이상일 때는 9 또는 10 정도로 제한하고 다른 bit를 조금씩 조절한다. 이렇게 모든 ji 에 대하여 b_{ji} 를 구하여, codebook size를 결정한다.

이상에서와 같이 각 class의 ac energy에 의하여 VQ될 모든 vector의 codebook size를 결정한 다음, 아래와 같은 과정으로 DCT-CVQ의 codebook을 design한다.

- (1) 먼저 codebook을 만들기 위해 선택된 영상을 8×8 block으로 나누어 각 block에 대하여 2D DCT한다.
- (2) 각 block을 그림 4와 같은 과정으로 6개의 class로 분류한다.
- (3) 변환계수들을 그림 5와 같이 몇 개의 vector \underline{V}_i 로 분할하여 해당하는 class의 training vector로 저장한다.
- (4) 각 class에 속한 vector \underline{V}_i 의 training vector를 이용하여, 각 class의 각 vector에 대해서 미리 정해진 codebook size로 각각 codebook을 design한다.

5. 실험 및 결과

4절에서 설명한 DCT-CVQ 방식을 다음과 같은 실험조건으로 computer simulation을 통하여 주관적 평가와 $10 \log_{10} [255^2 / \text{MSE}]$ 로 정의되는 SNR로 성능을 평가하였다.

class의 수 : 6

class 전송 : Huffman coding [9]

DC coefficient quantization : uniform scalar quantizer

VQ codebook design : LBG algorithm [1]

Training vector로 사용한 image : 512×512 3장,
256 \times 256 5장

Test image : Girl (training vector에 포함되지 않음)

기존의 방식과 비교하기 위하여, classification을 하지 않는 DCT-VQ방식과, Aizawa 등에 의하여 제안된 A-DCT-VQ 방식[7]에 대해서 simulation을 하였다. A-DCT-VQ방식은 ac energy에 의하여 4개의 class로 나눈다음, 각 class에 같은 수의 training vector가 포함되게 하여 VQ하는 방식이다.

그림 7은 0.470 bpp (30 bit/64 pel)에서 DCT-VQ, A-DCT-VQ 방식과 본 논문에서 제안한 방식인 DCT-CVQ 방식에 의하여 각각 simulation된 결과를 나타내는데, DCT-CVQ가 다른 DCT-VQ 방식에 비하여 edge에서의 degradation이 훨씬 적은 것을 알 수 있다. 그림 8은 0.343-0.547 bpp의 4가지 bit rate에서 각각 SNR를 비교한 결과를 나타낸다. 그 결과를 보면 DCT-CVQ방식이 가장 높은 결과를 나타냈는데, A-DCT-VQ 방식보다 대체로 1.5-3.0 dB 정도의 성능향상을 보였다.

6. 결 론

DCT domain에서의 2개의 변환계수에 의한 classification은 그 방법이 매우 간단하고, edge에 대한 DCT domain에서의 ac energy 분포의 특성을 잘 나타낸다. 여기에서 vertical, horizontal class는 energy 집중이 가장 강하고 diagonal class는 ac energy가 전체적으로 커지는 특성을 갖는다.

이와같은 classification을 한 다음 ac energy가 많은 class에 많은 bit를 할당해주는 adaptive coding 기법을 사용함으로써, 전체적으로 작은 비율을 차지하지만, 시각적 효과에 가장 큰 영향을 주는 edge를 정확하게 재생할 수 있기 때문에 0.5 bpp 미만의 낮은 bit rate에서 DCT-CVQ system은 다른 DCT-VQ 방식에 비하여 뛰어난 주관적 성능을 나타내었다. 그리고 이런 DCT-CVQ coding 방식은 영상의 통계특성에 거의 영향을 받지 않기 때문에, mismatch 현상이 다른 VQ 방법에 비하여 거의 없었다.

여기에서는 계산량과 memory 문제 때문에 ac energy가 큰 class의 vector에 대해서 가장 큰 codebook size를 1024개로 제한했는데, 만약 edge의 특성을 좀더 잘 나타낼 수 있도록 가장 큰 codebook size를 좀더 크게 하여 codebook을 design한다면 edge degradation이 거의 없는 system을 설계할 수 있을 것이다. 그런데 계산량과 memory의 제한에 의하여 가장 큰 codebook size가 이와같이 제한된다면, 본 논문에서 제안된 DCT-CVQ 방식은 0.5 bpp 미만의 낮은 bit rate에 적합한 coding 방식이라 할 수 있다.



(a)



(b)



(c)



(d)

그림 7. 0.470 bpp에서 재생된 image (Girl)
 (a) Original (b) DCT-VQ, SNR=29.68dB
 (c) A-DCT-VQ, SNR=30.27dB
 (d) DCT-CVQ, SNR=32.72dB

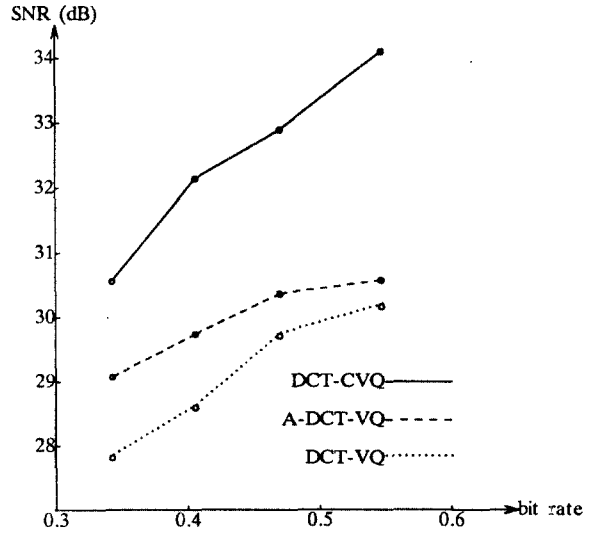


그림 8. SNR 비교

References

1. Y.Linde, A.Buzo, and R.M.Gray, "An algorithm for vector quantizer design," IEEE Trans. on comm., vol. COM-28, pp.84-95, Jan. 1980.
2. R.M.Gray, "Vec ∞ quantization," IEEE ASSP Magazine, vol.1, pp.4-29, Apr. 1984.
3. B.Ramamurthi and A. Gersho, "Classified vector quantization of images," IEEE Trans. on comm., vol. COM-34, No.11, pp.1105-1115, Nov. 1986.
4. C. Labit and J.P.Maresq, "Vector quantization in transformed image coding," ICASSP 86, pp.145-148.
5. T.Omachi, Y.Takashima, and H.Okada, "DCT-VQ coding scheme using categorization with adaptive band partition," pp.161-162, PCS 87.
6. W.H.Chen and C.H.Smith, "Adaptive coding of monochrome and color images," IEEE Trans. on comm., vol.COM-25, No.11, pp.1285-1292, Nov. 1977
7. K.Aizawa, et al., "Adaptive discrete cosine transform coding with vector quantization for color images," pp.985-988, ICASSP 86.
8. N.Ahmed, T.Natarjan and K.R.Rao, "Discrete Cosine Transform," IEEE Trans. on computers, pp.90-93, Jan. 1974.
9. D.A.Huffman, "A method for the construction of minimum redundancy codes," Proc. IRE vol.40, Sept. 1952.