

論 文

저 전송율에서의 영상 부호화를 위한
멀티테이블 벡터 양자화기

正會員 全 俊 鉉* 正會員 金 在 均*

Multi-table Vector Quantizer
for Image Coding at Low RatesJoon Hyeon JEON*, Jae Kyoon KIM* *Regular Members*

要 約 본 논문에서는 저 전송에 따른 벡터 양자화의 에지 결손(edge degradation)을 줄이기 위한 새로운 방법을 제안하였다. 이 방법은 여러개의 codebook들을 사용하며, 이 codebook들 중 하나는 현재 부호화할 벡터의 표준편차에 의해 선택되어진다. 또한 codebook을 찾는 계산량을 줄이기 위해 SD 탐색(search) 방법을 제안하였다. 결과로서 제안된 벡터 양자화기는 0.4-0.7 bpp 전송율에서 30dB 이상의 좋은 성능을 얻었다.

ABSTRACT The paper describes a new algorithm for reducing the edge degradation of a vector quantization(VQ) in low rates. This algorithm exploits several codebooks, one of which is adaptively selected by standard deviation(SD) of the vector currently being coded. Also, SD search method for reducing the complexity of codebook search is proposed. As a result, a proposed VQ has good performance about 30 dB PSNR in the 0.4-0.7 bpp range.

I. 서 론

VQ(Vector Quantizer)를 저 전송율을 이용한 부호화에 적용하기 위해서는 여러가지 문제점들이

있다. 그 중에서 에지 부근의 중요한 성능감소가 나타남을 알 수가 있다. 이러한 현상은 codebook의 설계시 training set에 사용되지 않은 통계특성이 다른 입력 영상의 경우 성능 저하가 더욱 심하게 일어난다. 그 결과 VQ의 좋은 성능을 위하여 벡터의 수를 늘리게 되면 전송량 및 입력 벡터에 해당되는 codevector를 찾아내는 필요한 계산량이 많아지게 된다.⁽¹⁾ 이와같이 1bpp(bit per pel) 미만의 저전송에서 VQ를 이용하여 영상을 부호화할

*韓國科學技術院 電氣및 電子工學科
Korea Advanced Institute of Science and Technology
Electrical Engineering.
論文番號 : 89-11(接受1988. 12. 10)

때 생기는 여러가지 문제점들 가운데 에지 결손 (edge degradation)과 codebook 탐색에 따른 계산량을 줄이기 위한 노력이 계속 되어왔다.^(1,2,3)

Ramamurthi 등은 저 전송율에 따른 에지 결손 방지를 억제하기 위하여 codebook을 에지 종류에 따라 partial codebook의 집합으로 분류 및 설계하여 입력 벡터의 에지 특성에 맞게 부호화하는 CVQ(Classified VQ)를 제안하였다.⁽²⁾ 이 방법은 계산량 및 에지 결손을 줄일 수가 있었으나 벡터 크기의 증가(4×4, 5×5)시 에지 종류가 급격히 증가하여 codebook 크기의 증가와 정확한 에지 분류가 어려운 단점이 있다. 또한 codebook 내의 각 codevector의 평균은 training 벡터에 종속하게 구하였기 때문에 재생 영상에서 경제현상 등의 중요한 성능 감소의 원인이 된다.⁽²⁾ Lee and Lee 가 제안한 GSVQ(Gain-Shape VQ)는 codevector로부터 평균과 gain을 분리하여 codebook을 설계하므로써 codebook의 통계적인 안전성을 높였다.⁽³⁾ 이 방법은 training set 밖의 영상에 대해서도 안정된 성능을 얻을 수가 있었으나 평균과 gain에 대한 추가 정보의 증가로 shape codebook 크기의 감소와 벡터의 평균과 gain의 양자화 오차에 대한 누적이 중요한 성능 감소의 원인이 된다. 결과로써 위에서 언급된 CVQ 및 GSVQ 방법들은 0.7bpp 이하에서는 만족할 만한 성능을 얻지 못했다.^(2,3)

제안된 MTVQ(Multi-table VQ)는 0.7bpp 이하의 저 전송율방식으로 벡터의 평균을 따로 전송하는 평균 분리(mean separation)와 여러개의 독립적인 codebook을 사용하는 multi-table 방식을 사용하였다. Multi-table 방식이란 training set의 통계 특성에 따라 분류된 에지 변화 정도에 따른 SD (Standard Deviation) 영역에 따라 각각 독립적인 codebook을 두는 방식으로 입력 벡터 activity (즉 SD)에 따라 해당하는 SD 영역의 codebook으로부터 codevector를 얻게 된다. 이 방법은 Chen 이 제안한 adaptive transform에서 블럭의 activity에 따라 bit map을 달리 선택하는 방식과 유사하다고 할 수가 있다.⁽⁴⁾ VQ의 다른 중요한 문제점들 중의 하나인 codebook 탐색에 있어서도 입력 벡터의 변화 범위에 따라 codebook의 탐색(search)

범위를 달리하는 SD 탐색을 본 논문에서 제안하였으며 재생 영상의 화질을 손상시키지 않는 범위에서 많은 계산량을 줄일 수가 있었다. 이와 같이 MTVQ는 입력 벡터의 지역 통계특성 변화를 충분히 고려하여 벡터들의 각 SD 영역에서의 발생 확율을 근거로한 부호화 방식으로 각 SD 영역의 특성에 맞게 효율적인 비트(즉 codebook)를 할당하기 때문에 에지 결손 방지와 outside training vector sequence의 입력 영상에 대해서도 높은 성능을 기대할 수가 있다. 또한 training vector sequence의 확률에 근거한 최적 codebook 설계와 최적 스칼라 양자화기(scalar quantizer)의 설계로 성능을 향상시켰으며, DPCM을 통한 블럭의 평균 전송과 Huffman coding⁽⁵⁾ 방식을 도입하여 효율적인 부호화를 하였다.

제2장과 제3장에서는 MTVQ의 기본 방법 및 최적 설계에 대해서 각각 서술하였으며, 제4장에서는 MTVQ의 성능 평가와 GSVQ와의 성능 비교를 하였다. 끝으로 제5장에서는 결론을 맺었다.

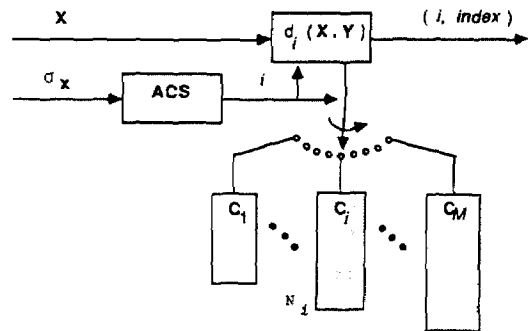
II. Multi-table Vector Quantizer(MTVQ)

본 논문에서는 벡터의 평균 분리와 multi-table 방식을 사용하므로써 평균과 SD의 통계 특성 변화에 따른 codebook의 안전성을 높였으며, 확률에 근거한 부호화 방식을 택하므로써 0.7bpp 이하의 저 전송율에서도 성능이 좋은 MTVQ를 제안하였다. 그림1은 multi-table 방식을 사용한 MTVQ의 적응식 codebook 선택(adaptive codebook selection)을 나타낸 것으로 MTVQ의 codebook은 class로 구분되는 각 SD 영역마다 독립적인 고유의 codebook으로 구성되어 있다. 따라서 입력 벡터는 SD에 따라 ACS(Adaptive Codebook Selector)로부터 class를 선택하여 해당하는 codebook을 사용하여 전송될 벡터를 찾는다. 그림1에서 C_i 와 N_i 는 M개의 class 중 i번째 class의 codebook과 codebook 크기를 나타낸 것이며 i의 값이 클 수록 SD 값, 즉 에지의 변화가 큰 class를 나타낸다. 따라서 송신측은 벡터마다 평균에

대한 정보 및 class 정보 i 와 해당하는 codebook에서 찾은 재생 벡터의 index를 수신측에 전송해야 한다. 그림2는 표1과 같은 통계특성을 갖는 512×480 크기의 2가지 흑백영상들을 4×4 크기를 갖는 벡터들의 집합으로 하였을때 각 벡터들의 SD에 따른 누적 확률분포도를 나타낸 것이다. 표2는 이러한 확률분포를 갖는 벡터들에 대해 SD 변위 범위에 따라 2,3,4,5개의 영역으로 나누었을 때의 확률을 나타낸 것으로 각 영역과 중심 SD값은 최적 스칼라 양자화기(scalar quantizer) 설계방식인 Lloyd algorithm⁽⁶⁾을 사용하여 구하였다.(Appendix A 참조) 이 표에서 각 class에 속한 벡터들의 평균 분리에 따른 평균 MSE(Mean Squared Error)를 구한 결과 class에서 왜곡(distortion)이 큰 것을 알 수가 있다. 다시 말해 변화가 심한 에지를 갖는 벡터들에 대한 고려가 필요하다. 이와 같이 SD에 따라 영역을 나눌 경우 벡터들의 에지 변화 특성에 맞게 codebook을 따로 설계할 수 있을 뿐만 아니라, 변화율에 따른 평균 왜곡을 계산하여 각 class의 codebook 크기를 좀더 효율적으로 최적화할 수가 있다. 즉 변화가 심한 class들의 codebook 크기를 크게 하므로써 에지 결손을 막고 변화가 작은 class들의 codebook 크기를 줄이므로써 효과적인 비트 할당에 의한 부호화 효율을 높일 수가 있다. 또한 실제 부호화에 있어서 재생영상에 대해 시각효과(human visibility)를 향상시키는 결과를 가져오게 되었다.

실제 VQ에서 codebook 설계시 일반적으로 많이 사용하는 MSE(Mean Squared Error)는 codebook 설계에서 에지를 포함한 벡터들의 왜곡 측정(distortion measure)에는 적합하지 않는데, 이것은 training set에는 전체 블럭 수에 비해 에지 블럭들이 차지하는 비율이 작기 때문에 결국 codebook 설계에 따른 전체 MSE에서 중요하게 고려되지를 못한다. 이것은 재생 영상에서의 에지 결손의 원인이 되며, 이것을 방지하기 위한 다른 왜곡 측정이 필요하다. 따라서 codebook을 설계할 때 각 class의 에지 변화 특성에 맞게 효율적인 codebook 크기를 할당하기 위해 각 class에 속하는 벡터들의 왜곡을 계산할 때 각 class의 SD 변화

특성에 따라 왜곡의 가중치를 달리하는 WMSE(Weight Mean Squared Error)를 사용하였다. (Appendix B참조)



- X** : Mean separated input vector
- Y** : Codevector
- C_i** : Codebook of the i th class
- ACS** : Adaptive Codebook Selector
- $d_i(\dots)$** : Distortion measure
- σ_x** : Standard deviation of input vector

그림 1. MTVQ에서 입력 벡터의 표준편차를 사용한 적응식 codebook 선택
Adaptive codebook selection of MTVQ according to standard deviation of input vector.

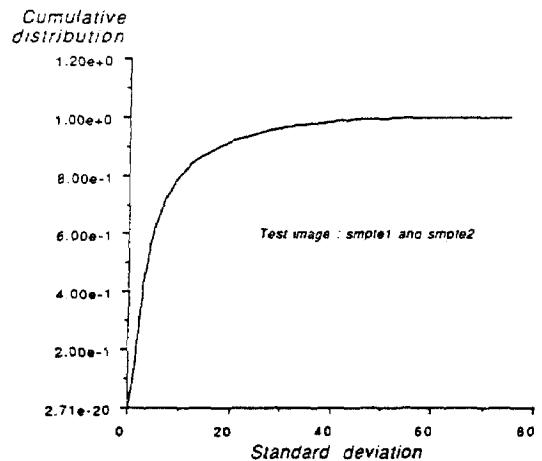


그림 2. 표준편차에 따른 실험 영상들의 누적확률 분포
Probability distribution versus standard deviation

표 1. 실험 영상들의 통계 특성
Stochastic characteristics of test image

	SMPTE1	SMPTE2
Mean	-32.74	-29.76
Standard deviation	45.94	46.58

표 2. 각 SD 영역에 속한 벡터들의 확률
Probability of vectors belong to each SD region.

Number of class	Centered SD values	Probability	Average RMSE per vector
2	5	0.838	4.3
	20	0.162	24.0
3	3	0.716	3.4
	11	0.207	12.4
	31	0.077	32.6
4	3	0.616	2.9
	8	0.205	7.9
	15	0.118	16.6
	34	0.061	35.1
5	1	0.074	1.1
	2	0.358	2.5
	5	0.345	5.7
	13	0.156	14.6
	33	0.067	34.2

III. MTVQ 설계

Multi-table 방식에서 codebook을 선택하는 방법으로 벡터의 SD(Standard Deviation)을 사용하는 이 방법은 먼저 다음 세가지가 최적화되어야 한다.

- i) DPCM내에서의 스칼라 양자화기
- ii) 적응식 codebook 선택기

iii) 각 class의 codebook 크기 할당

위 세가지를 만족하는 MTVQ 부호기가 설계되었을 경우 한개의 k 차원 입력 벡터를 전송하는데 필요한 평균 비트 B_{av} 는 벡터의 평균, class 및 codebook에 따른 평균 전송비트를 각각 B_{μ}, B_c, B_v 라 하면 식(1)과 같다.

$$B_{av} = B_{\mu} + B_c + B_v \quad (\text{bit/vector}) \quad (1)$$

만약 codebook이 M개이고, 임의의 입력 벡터가 i번째 class 영역에 속한 확률을 P_i , i번째 class의 codebook 크기를 N_i , 재생 벡터의 전송 비트를 $B_i = \log_2 N_i$ 라고 하면 식(1)의 B_v 는 다음과 같다.

$$B_v = \sum_{i=1}^M P_i B_i = \sum_{i=1}^M P_i \log_2 N_i \quad (2)$$

식 (1)과 (2)는 위의 세가지 최적화 조건을 만족하는 부호화기를 설계하는데 매우 중요하다. 먼저 평균 전송율(bit / vector) B_{av} 가 정해지고 최적의 스칼라 양자화기가 설계되면 k 차원의 training vector set에 대해 블럭 평균의 전송 평균 비트 B_{μ} 를 구한후 class의 수 M과 각 codebook 크기를 최적화 시켜야 한다.(Appendix A와 appendix B 참조)

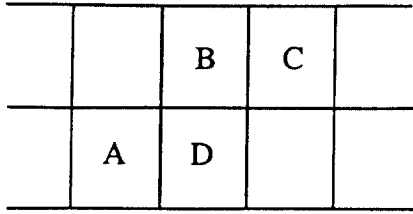
A. 최적 스칼라 양자화기 설계

본 논문에서는 벡터의 스칼라 평균에 대한 정보를 최소화하면서 좋은 성능을 유지하기 위해 DPCM(Differential Pulse Code Modulation) 방식과 Lloyd가 제안한 최적 비균등 양자화기⁽⁶⁾을 사용하였으며, 예측 오차(prediction error)의 통계 특성에 근거한 Huffman coding을 하였다.

k 차원 입력 벡터 $\mathbf{X} = \{x_1, x_2, \dots, x_k\}$ 의 평균을 $\mu = 1/k \cdot \sum_{i=1}^k x_i$ 라하고 양자화 된 값을 $\mu(\cdot)$, DPCM 재생 값을 $\mu(\cdot)$ 라 할 때 그림3과 같은 예측 방법을 사용하였을 경우 현재 블럭의 예측 값 μ_p 는 식 (3)과 같다. 이때 예측값은 실험적으로 얻어졌다.

$$\mu_p = 1/8 \cdot (3\hat{\mu}_A + 3\hat{\mu}_B + 2\hat{\mu}_C) \quad (3)$$

위와 같이 4×4 크기의 벡터 평균에 대해 DPCM 한 결과 MTVQ의 평균 전송율이 0.7 bpp 이하일 때 9 level의 비균등 양자화기 사용이 적합한 것으로 실험결과 나타났다.



D : Current block
A, B, C : Reference block

그림 3 예측을 위한 블록의 위치
Block allocation for prediction.

B. 최적 ACS(Adaptive Codebook Selector) 설계

만약 k 차원 입력 벡터 $X = \{x_1, \dots, x_k\}$ 의 평균과 SD를 각각 μ_x 와 σ_x , ACS에 의해 선택될 class의 SD 영역 집합을 $A = \{A_i | i=1, \dots, M\}$ 라 하고, 이때 i번째 class의 SD 영역 A_i 의 대표 SD를 σ_i 라고 가정할 경우 만약 입력 벡터 X의 SD 값이 ACS에 의해 α 번째 class의 SD 영역을 만족하는 왜곡(distortion) $d(\sigma_x, \sigma_\alpha)$ 을 최소화 한다면 ACS는 최적이라고 할 수가 있다. 이 최적의 ACS가 되기 위해서는 다음 두가지 조건을 만족해야 하는데, 이 조건들은 Lloyd가 MSE를 사용하여 설계한 최적 quantizer 성질과 같다.

Property 1) $\min_{i \in M} d(\sigma_x, \sigma_i) = \min_{\sigma_\alpha \in A_\alpha} d(\sigma_x, \sigma_\alpha) \quad (4)$

Property 2) $\text{Cent}(A_\alpha) = \sigma_\alpha = \frac{1}{i(\alpha)} \sum_{\sigma_{x_i} \in A_\alpha} \sigma_{x_i} \quad (5)$

여기서 식 (5)는 α 번째 class 영역의 무게 중심(centroid) 값이며, $i(\alpha)$ 는 A_α 에 속하는 입력 벡터의 수이다. 여기서 사용된 벡터 및 class의 모든 SD 값은 반올림한 정수값을 취하여 사용하였으

며, class에 대한 정보는 Huffman coding을 하였다.

C. 최적 codebook 설계

Training input vector set에 대해 벡터 평균의 전송 평균 비트 B_μ 와 appendix A에 따른 class의 수 및 codebook의 평균 전송비트 B_c 와 B_v 가 결정되면 각 codebook 크기가 최적화 된다. 각 class가 최적의 codebook 크기 $\{N_i^* | i = 1, \dots, M\}$ 을 가질 조건은 식 (6)과 같다. (Appendix B 참조)

$$D_i^* = \text{constant for all class} \quad (6)$$

여기서 D_i^* 는 i번째 class의 최소 왜곡(minimum distortion) 값이며, 각 class의 최소 왜곡이 서로 같을 때 각 class의 codebook 크기는 최적이라 말할 수 있다. 최적조건 식 (6)으로부터 i번째 class에 할당된 최적의 codebook 크기 N_i^* 는 주어진 B_v 에 대해서 식 (7)과 같이 얻어진다.

$$\log_2 N_i^* = B_v + \log_2 [K_i^{1/\beta} \cdot \sigma_i] - 0.5 \sum_{i=1}^M P_i \log_2 [K_i \cdot \alpha^\beta] \quad (7)$$

여기서 P_i 는 k 차원 입력 벡터가 i번째 class일 확률이며, β 는 $\frac{k}{2}$ 이다. σ_i 는 i번째 class의 중심 SD 값으로써, 각 class의 activity에 따라 codebook 크기를 효율적으로 할당하기 위한 가중치이다. 또한 K_i 는 i번째 class의 joint probability를 고려한 상수로써 appendix B에서 구해진다.

식 (7)은 Yamada⁽⁸⁾ 등이 Gersho⁽⁹⁾가 구한 왜곡 D의 경계(bound)에서 양자화계수를 최적화하여 구한, 최적의 경우 경계 왜곡 D^* 를 사용하여 얻은 것이다. 또한 최적 codebook 크기 집합 $\{N_i^*\}$ 를 더욱 효율적으로 결정하기 위해서 appendix B의 식 (B-2)의 WMSE를 사용하였으며, 본 논문에서는 각 class에 해당하는 중심 SD값의 제곱을 가중치로 사용하였다. 이것은 예지 결손이 큰 class에 크기가 큰 codebook을 할당하므로써 성능

향상을 가져오도록 하였다. 식 (7)로부터 각 class의 최적 codebook 크기가 할당되면 각 class에 해당하는 k 차원 training 벡터 집합에 대하여 LBG algorithm을 사용하여 각 codebook을 최적화한다. 이때 사용되는 왜곡 측정은 일반적으로 많이 사용되는 MSE를 사용하였으며, 각 SD 영역의 경계(boundary)는 이웃하는 각 class들이 공유하도록 중첩(overlap)하여 codebook을 설계하였다.

D. MTVQ Coder

0.7 bpp 이하의 저 전송 부호화를 위한 성능이 좋은 부호화기를 설계하는 것은 매우 어렵다. 특히 매번의 입력 벡터 마다 평균 분리에 따른 평균 값과 class에 대한 부가 정보를 전송한다는 것은 전송율면에서 한계가 따른다. 그림4는 이러한 문제점을 해결하기 위해 제안한 MTVQ 방식의 송신단과 수신단의 블럭도이다. 우선 K 차원 입력 벡터 $X = \{x_1, x_2, \dots, x_k\}$ 에 대해 평균과 SD를 계산하여 평균은 DPCM 하여 수신측에 전송하고 SD는 codebook 선택기(selector)인 ACS에 의해 codebook을 지정하는데 사용된다. 이때 MTVQ는 정해진 codebook으로부터 inverse DPCM에 의해 재생된 평균 $\hat{\mu}$ 에 의해 평균 분리된 벡터 \hat{X} 에 대한 재생 벡터를 찾아 해당하는 index를 수신측에 전송하게 된다. DPCM내의 양자화기는 앞에서 설명한 Lloyd의 최적 스칼라 양자화기를 사용하였으며, 양자화된 평균 값과 지정하는 class에 대한 정보는 Huffmam coding을 하여 전송율을 최소화하였다. 표3은 앞에서 제시한 최적 조건들을 바탕으로 표2의 실험영상들을 training 벡터 집합(4x4 크기)로 사용하였을때 평균 전송율 0.4 bpp-0.7 bpp에 따라 최적화된 class의 수와 각 class의 codebook 크기를 나타낸다. 이때 평균에 대한 정보는 모두 9 레벨(level)을 사용하였다.

E. SD(Standard Deviation) 탐색(search)

VQ에서 중요한 문제 중의 하나는 codebook에서 입력 벡터와 최소 왜곡을 갖는 codevector의 탐색에 따른 많은 계산량이다. 이러한 code-

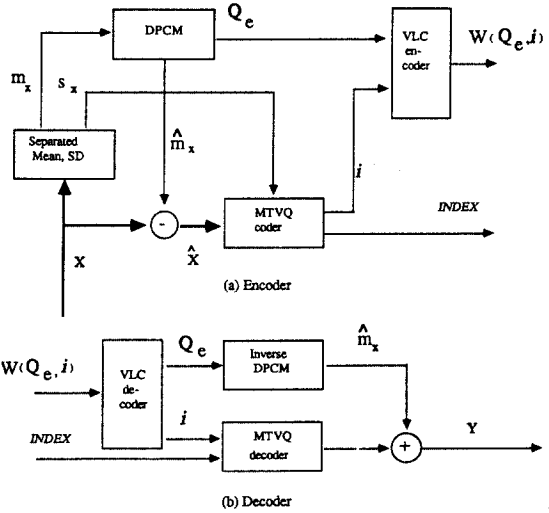


그림 4. MTVQ 부호화기
MTVQ coder

표 3. 주어진 B_{av} 에 대해 최적화된 class의 수와 각 codebook size
optimized class number and code book sizes of each class for a given B_{av} .

B_{av}	M	N_1	N_2	N_3	N_4	N_5
5.5	2	0	2048			
6.5	3	0	1024	2048		
8	4	0	512	1024	2048	
11	5	0	128	512	1024	2048

book 탐색에 따른 많은 계산량을 줄이기 위해 그동안 여러가지 탐색 방법들이 제안되었다.⁽¹⁾ 제안된 방법들은 codebook의 일부분만을 탐색하는 부분 탐색으로서 계산량 및 시간은 절약할 수는 있으나 codebook 크기 전체를 탐색하는 완전 탐색(exhaustive search) 보다는 PSNR(peak Signal-to-Noise ratio) 및 재생 영상의 화질면에서 성능 감소가 불가피하다.⁽¹⁾ 본 논문에서는 이와같은 부분 탐색의 한 종류로서 매 입력 벡터마다 효과적으로 codebook 탐색 범위를 달리하여 재생 영상의 주관적 화질(subjective quality)에 영향을 미치지 않으면서 계산량을 줄인 SD search를 제안하였다.

SD search는 선택된 codebook으로부터 입력 벡터의 SD 값을 사용하여 탐색할 SD 영역을 구하여 그 영역에 해당하는 부분 codebook만을 탐색하는 방법이다. 따라서 MTVQ의 ACS는 매번 입력 벡터의 SD를 사용하여 부호화할 codebook을 선택하는 동시에 SD 탐색에 따른 탐색 SD 영역을 구하여 부분 codebook의 범위를 부호화기에 지정해 주어야 한다. 식 (8)은 입력 벡터로부터 얻어지는 탐색할 SD 영역이다.

$$[(1-\alpha) \cdot \sigma_x] \leq \text{search SD region} \leq [(1+\alpha) \cdot \sigma_x] \quad (8)$$

여기서 bracket[·]는 버림(truncation) 정수의 값을 의미하며, σ_x 는 입력 벡터의 SD이다. 또한 α 는 탐색 영역을 결정하는 상수로서, 실험 결과 $\alpha=0.15$ 일때 PSNR과 재생 영상의 주관적 화질면에서 만족할 만한 결과를 얻었다. 이때 SD 탐색은 완전 탐색(exhaustive search)에 비해 약 90%의 계산량을 줄일 수가 있었다. 그림5는 $\alpha=0.15$ 일때 입력 벡터의 SD에 의해 구해진 i번째 class의 codebook에서의 탐색 SD 영역과 해당하는 부분 codebook을 보여준다.

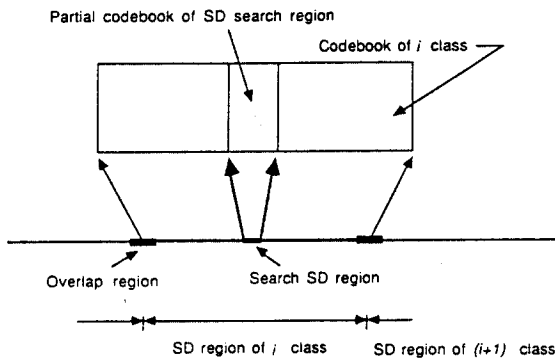


그림 5. SD search
SD(standard deviation)starch.

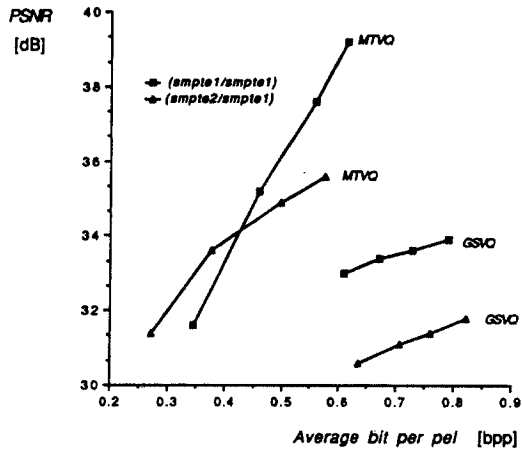
IV. 시뮬레이션(Simulation) 결과 및 검토

먼저 본 논문에서 제시하는 MTVQ 부호기의 성능을 평가하고 비교하기 위해 표1의 통계특성을 갖는 두가지 영상, SMPTE1과 SMPTE2을 Data Base 및 training set로 사용하여 PSNR(즉, $PSNR = 10\log_{10}(255^2/MSE)$)을 구했으며 Lee and Lee가 제안한 GSVQ⁹⁾와 비교하였다. 이때 사용된 벡터 크기는 4×4 로 고정하였다. 그리고 입력 벡터마다 탐색 SD 영역을 구하여 해당하는 부분의 codebook만을 탐색하는 SD 탐색 때의 PSNR 및 재생 영상의 주관적 실험(subjective test)을 하였다. SD 탐색에 따르는 PSNR 및 주관적 실험은 일반성을 주기 위해 표1과 통계특성이 다른 256×256 크기의 8개의 흑백영상을 training vector set(4×4 크기)로 사용하였을 때 설계된 최적의 MTVQ 부호기에 표1의 2가지 영상을 적용하여 행하였다.

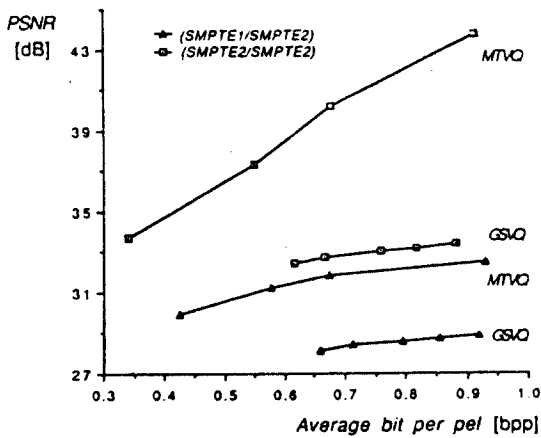
A. GSVQ와의 성능 비교 및 검토

제안된 MTVQ와 GSVQ와의 성능 평가는 표 1의 2가지 실험 영상들에 대해 부호화 영상과 codebook 설계 영상을 조합으로 하는 4가지 방법으로 비교하였으며 그림6와 같다. 예를 들면 그림 6.(a)는 부호화 영상을 각각 SMPTE1과 SMPTE2 로 하고, codebook 설계 영상을 SMPTE1 을 사용하였을때 성능비교를 한 것으로서 표기는 (SMPTE1 / SMPTE1)과 (SMPTE2 / SMPTE 1)로 나타내었다. 그림6에서 나타난 바와 같이 MTVQ는 평균 0.4-0.7 bpp의 낮은 전송율에서 영상의 종류에 관계없이 30 dB 이상의 PSNR의 높은 성능을 갖는다는 것이 증명되었다. 또한 GSVQ 는 평균 전송 비트의 증가에 비해 미미한 PSNR 증가를 나타내는 반면 MTVQ는 비트 증가에 따른 PSNR 증가가 뚜렷하다는 것을 알 수가 있다. 두 가지 방법에 대해 재생영상들을 비교한 결과 0.7bpp 이하에서의 MTVQ의 화질이 매우 우수함을 알 수가 있었으며, 특히 GSVQ는 gain에 대한 오차 누적이 0.8 bpp 미만의 재생 영상의

화질에 상당히 많은 영향을 끼쳐 에지가 오히려 뭉개지거나 불분명해지는 현상을 볼 수가 있었다. 이때 사용된 GSVQ의 평균과 gain은 각각 8 레벨과 9 레벨의 DPCM과 Huffman coding을 하였다.



(a)



(b)

그림 6 MTVQ와 GSVQ의 성능비교.
Performance comparison to MTVQ with GSVQ.

B. SD 탐색(search)에 따른 주관적 실험

본 논문에서는 부분 탐색에 따른 성능 감소를 줄이기 위해 입력 벡터의 변화특성에 따라 탐색 범위를 달리하는 SD 탐색을 제안한 바 있다. 그림 7은 MTVQ의 성능 평가에, 좀더 일반성을 주기 위해 codebook 설계시 10개의 흑백 영상을 사용하여 통계 특성이 다른 표1의 영상들을 적용하였을 때 평균 전송율에 따른 SD 탐색과 완전 탐색(exhaustive search)의 성능 비교를 나타낸 것이며, 그림8과 그림9는 SD search를 했을 때 평균 전송율에 따른 재생 영상들을 보여주는 것이다. 이때 탐색 SD 영역은 $\alpha=0.15$ 일때 식 (8)에 따른다. 결과로서 SD 탐색은 완전탐색에 비해 전체 codevector 수의 8%-10% 만을 탐색함에도 불구하고 성능 감소가 그리 심하지 않았음을 알 수가 있으며, 그림8과 그림9의 재생 영상들에 있어서 전송율에 비해 화질이 우수함을 알 수가 있다. 그림8과 그림9의 재생영상들에서 전송율이 0.5 bpp 미만에서 에지 결손에 따른 성능은 향상시켰으나, 동일 영역(homogeneous region)의 경우 경계 현상들이 나타난 것을 볼 수가 있다. 이것은 낮은 activity(첫번째 class)에 속하는 벡터들에 대해서는 평균만을 전송한다는 점과, 또한 동일 영역(homogeneous region)에 잡음(white noise)

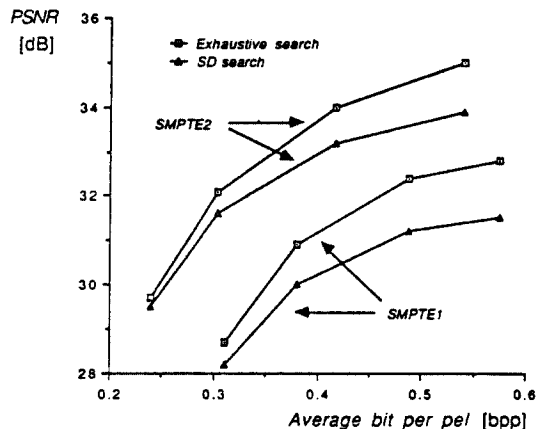


그림 7. SD search와 exhaustive search와의 성능 비교
Performance comparison of SD search with exhaustive search



(a) Original



(b) 0.31bpp, PSNR=28.2dB



(c) 0.38bpp, PSNR=30dB



(d) 0.49bpp, PSNR=31.2dB

그림 8 SD search에 따른 재생 영상(SMPTE 1)
Reconstructed image(SMPTE1) according to SD search.



(a) Original



(b) 0.30bpp, PSNR=31.6dB



(c) 0.42bpp, PSNR=33.2dB



(d) 0.54bpp, PSNR=33.9dB

그림 9 SD search에 따른 재생 영상(SMPTE2)
Reconstructed image(SMPTE2) according to SD search.

이 섞일 경우 그 블록의 SD 값이 커져 발생하는 잘못된 codebook 탐색 때문이다. 평균만을 전송하는데 발생하는 경계 현상은 첫번째 class를 두개의 class로 나누었을 때 거의 제거할 수가 있는데, 그림 9.(d)의 재생 영상과 같다. 그러나 실제 부호화시 연속된 영상에서는 동일 영역의 경계 현상(0.5 bps 미만)은 잘 눈에 띄지 않는 시각적인 효과를 얻을 수가 있었다.

V. 결 론

본 논문에서 제안된 새로운 MTVQ 방식은 0.7 bps 이하의 저 전송 부호화 방식으로 평균 분리와 multi-table 방식을 사용하므로써 codebook의 통계특성에 따른 불안전성을 제거하였으며, 성능 감소가 적은 SD 탐색을 제안하므로써 많은 계산량을 줄였다. 그 결과 에지 결손(edge degradation)을 방지하여 시각적(human visibility) 효과에 따른 성능향상을 가져 왔다. 또한 최적 ACS(Adaptive Codebook Selector) 설계, 평균 분리에 따른 최적 스칼라 양자화기 설계, 그리고 최적 codebook 설계에 의해 부호화기의 효율을 최대로 하였으며, Huffman coding을 하

로써 비트의 증가를 억제하였다. 제안된 이 MTVQ의 장점은 입력 벡터마다 평균 분리와 SD에 의한 입력 특성에 맞는 codebook을 선택하기 때문에 outside training sequence의 영상에 대해서도 좋은 성능을 나타낼 수가 있다는 점이다.

Appendix A Optimal MTVQ design

Training vector sequence들에 대해 전송율(bit / vector)이 정해지면 최소왜곡(minimum distortion)을 갖는 최적 class의 수와 각 class의 최적 SD 영역 및 최적 Codebook을 설계하는 것이 중요하다. 식 (1)과 식 (2)로부터 MTVQ의 최적 ACS(Adaptive Codebook Selector)와 최적 codebook의 설계과정은 다음과 같다.

(1) Initialize the class number to $M=2$ and the distortion to $D_1=D_0$. Also, determine B_μ for block means using DPCM and huffman coder.⁽⁶⁾

(2) Determine the centered SD set $\{\sigma_i\}$ in each class by using the Lloyd algorithm, and obtain

B_c by using huffman coding. Therefore, $B_v = B_{av} - (B_u + B_c)$.

(3) Allocate the optimal set $\{\log_2 N_i\}$ based on eq.(7) for a given value of B_v below the condition of eq.(6), and design the optimal codebook set $\{C_i\}$ by using the clustering algorithm called the LBG algorithm.⁽⁷⁾

(4) Calculate the overall distortion D_M of eq.(3). If D_M is smaller than D_{M-1} , go to step(2) after $M=M+1$, D_M+D_{M-1} , otherwise stop.

Appendix B

Optimum codebook size assignment

X_i 를 i th class에 속하는 임의의 k 차원 입력 벡터라고 할때, 벡터 $X_i = \{x_{i1}, x_{i2}, \dots, x_{ik}\}$ 는 R_k^+ (i 번째 class에 속하는 k -dimensional Euclidean space)에서 i 번째 class에 속하는 벡터들의 joint probability density function, $P_{x_i}(x)$ 의 확률 분포를 갖는다. Q_i 를 codebook 크기가 N_i 인 i 번째 class의 k 차원 벡터 양자화기라고 가정하면, Q_i 의 재생 벡터의 집합 $\{Y_{i1}, Y_{i2}, \dots, Y_{iN_i}\}$ 은 식 (B-1)과 같이 R_k^+ 부분 영역 $\{S_{i1}, S_{i2}, \dots, S_{iN_i}\}$ 을 갖는다.

$$Q_i(X_i) = Y_{im}, \text{ if } X_i \in S_{im} \quad (B-1)$$

i 번째 class의 벡터 양자화기에서 왜곡 측정(distortion measure)으로 식 (B-2)와 같이 WMSE (Weighted Mean Squared Error)를 사용하면 보다 효과적인 codebook 크기를 결정할 수가 있다.

$$d_i(X_i, Y_{im}) = |(X_i - Y_{ik}) \cdot W_i \cdot (X_i - Y_{ik})^t|^{1/2} \quad (B-2)$$

식 (B-2)에서 $W_i = W_i I_i$ (W_i 는 scalar)이며 t 는 transpose을 나타낸다. 본 논문에서 가중치 (W_i)는 appendix A에서 구한 각 class의 중심 SD 값의 제곱을 사용하였다.(식 B-3)

$$W_i = \alpha_i^2 \quad i = 1, 2, \dots, M \quad (B-3)$$

위로부터 i 번째 class의 벡터 양자화에 의한 왜곡은 식 (B-4)와 같다.

$$D_i = \frac{1}{k} \sum_{m=1}^{N_i} \int_{S_{im}} d(X_i, Y_{im}) P_{x_i}(x) dx \quad (B-4)$$

여기서 S_{im} 는 i th class의 k 차원 공간에서 X_i 가 Y_{im} 로 대응하는 영역, 즉 $S_{im} = Q_i^{-1}(Y_{im})$ 이고 여기서 i 번째 class의 왜곡은 k 차원으로 정규화(normalization)된 평균 왜곡을 나타낸다.

Yamada⁽⁸⁾ 등은 Gersho⁽⁹⁾가 nonspherical region인 convex polytope와 양자화 계수를 도입하여 식 (B-5)와 같은 최적인 경우의 최소 경제 왜곡(minimum bound distortion) D_i^* 를 얻었다는 것을 보였다.

$$D_i^* = N_i^{-2/k} \frac{k \cdot \pi^{-1}}{k+2} \left\{ \left(\frac{k}{2} \right) \cdot \Gamma \left(\frac{k}{2} \right) \cdot (\det W_i)^{1/2} \right\}^{2/k} \cdot \|P_{x_i}(x)\|_{\frac{k}{k+2}} \quad (B-5)$$

여기서 왜곡 측정(distortion measure)으로 식 (B-2)의 WMSE를 사용했으며,

$P_i(X_i)$ 의 L_v norm은 식(B-6)과 같이 정의된다.

$$\|P_{x_i}(x)\|_v = \left[\int_{S_{im}} [P_{x_i}(x)]^v dx \right]^{1/v} \quad (B-6)$$

Appendix A의 최적 ACS 설계에 의해 class의 수 M 이 정해진면, 이때 전체 왜곡은

$$D = \sum_{i=1}^M P_i D_i^* = \sum_{i=1}^M \frac{P_i K_i (\det W_i)^{\beta/2}}{N_i^\beta} \quad (B-7)$$

이며, 여기서 N_i 는 i th codebook의 크기이고, $\beta = 2/k$, $W_i = W_i I_i$ 이다. 또한 K_i 는

$$K_i = \frac{k \cdot \pi^{-1}}{k+2} \left\{ \left(\frac{k}{2} \right) \cdot \Gamma \left(\frac{k}{2} \right) \right\}^\beta \cdot \|P_{x_i}(x)\|_{\frac{k}{k+2}} \quad (B-8)$$

이며, P_i 는 i th class에 속하는 벡터들의 확률이다. i th class의 최적 codebook 크기 N_i^* 는 식(B-7)을 최소화 하므로써 얻어진다. 따라서 식(B-7)의 D 를 최소화하기 위해 standard Lagrange multipliers method를 사용하면 식(2)로 부터 식(B-9)와 같으며,

$$D = \sum_{i=1}^M \frac{P_i K_i [\det W_i]^{\beta/2}}{N_i^\beta} + \lambda \left(P_i \sum_{i=1}^M \log_2 N_i - B_v \right) \quad (B-9)$$

각 class의 $\{N_i | i = 1, 2, \dots, M\}$ 에 관하여 D 를 미분하면

$$\lambda = \frac{\beta K_i [\det W_i]^{\beta/2}}{\alpha N_i^{\beta+1}}, \quad i = 1, 2, \dots, M \quad (B-10)$$

이 된다. 여기서 $\alpha = \log 2$ 이다. 식(B-10)에 식(B-7)의 $D_i^* = K_i [\det W_i]^{\beta/2} / N_i^{\beta+1}$ 을 대입하면

$$D_i^* = \frac{\alpha \lambda}{\beta} = \text{constant} \quad (B-11)$$

을 얻는다. 식(B-11)이 식(B-7)의 D 가 최소 왜곡(minimum distortion)을 갖을 때의 조건이 되며, 이것을 각 class의 최소 왜곡이 같을 때 최적 조건을 얻는다는 것을 말한다. 이때 식(2), 식(B-3)과 함께 최적 codebook 크기 N_i^* 를 구하면 식(B-12)와 같다.

$$\log_2 N_i^* = B_v + \log_2 [K_i^{\beta/2} \cdot \sigma_i] - 0.5 \sum_{i=1}^M P_i \log_2 [K_i^{\beta/2} \cdot \sigma_i] \quad (B-12)$$

參 考 文 獻

1. R.M. Gray, "Vector quantization, "IEEE ASSP Magazine, vol. 1, pp. 4-29, Apr. 1984.
2. B. Ramamurthi and A. Gersho, "Classified vector quantization of image", IEEE Trans. on comm., vol. COM-34, No. 11, pp. 1105-1115, Nov. 1986.
3. H.J. Lee and D.T.L. Lee, "A gain / shape vector quantizer for image coding", ICASSP 86, pp. 141-144.
4. W.H. Chen and C.H. Smith, "Adaptive coding of monochrome and color images", IEEE Trans. Comm., vol. COM-25, No. 11, pp. 1285-1292, Nov. 1977.
5. H. Murakami, S. Matsumoto, H. Yamamoto, "Algorithm for construction of Variable Length Code with limited maximum word length", IEEE TR. COMM., vol. COM-32, No. 10, pp. 1157-1159, October 1984.
6. S.P. Lloyd, "Least squared quantization in PCM, "Bell Lab Tech. Paper.
7. Y. Lind, A. Buzo, and R.M. Gray, "An algorithm for vector quantizer design", IEEE Trans. on comm., vol. COM-28, pp. 94-95, Jan. 1980.
8. Y.Y. Yamada, S. Tazaki, and R.M. Gray, "Asymptotic performance of block quantizers with difference distortion measures", IEEE Trans. Inform. Theory, vol. IT-26., No. 1, pp. 6-14, Jan. 1980.
9. A. Gersho, "Asymptotically optimal block quantization, "IEEE Trans. Inform. Theory, vol. IT-25, No. 4, pp. 373-380, July 1979.



全俊鉉(oon Hyeon JEON) 正會員
1961年8月19日生
1984年2月：東國大學校 電子工學科 卒業(工學學士)
1986年2月：韓國科學技術院 電氣 및 電子工學科 卒業(工學碩士)
1986年3月～現在：韓國科學技術院 電氣 및 電子工學科 博士課程



金在均(Jae Kyoong KIM) 正會員
1938年9月17日生
1958. 4～1962. 2：韓國航空大學 應用電子科(工學士)
1962. 3～1967. 2：서울大學校 大學院 電子工學科(工學碩士)
1967. 9～1971. 8：美國 남가주大學校大學院 電氣工學科(工學博士)
1962. 4～1966. 7：空軍服務 中尉
1967. 9～1972. 3：美國 남가주 大學校(研究院)
1972. 4～1973. 3：美國 NASA GSFC(研究院)
1973. 4～現在：韓國科學技術院院 教授
1984. 1～1985. 6：科學技術處 電氣·電子조정관