# Optimum design of two-dimensional subband filter banks using vector quantizer

Jonghong Shin*, Innho Jee**
*Department of Electrical Engineering, Hongik University
**School of Electronics, Electrical, and Computer Engineering, Hongik University

# 벡터양자기를 사용한 최적의 이차원 부대역필터의 구현

신 종 홍*, 지인호**
*홍익대학교 전기공학과, **홍익대학교 전자전기컴퓨터공학부

## Abstract

This paper provides a heuristic theory for modeling and analysis of vector quantization effects in 2-dimensional subband filter banks. This model is used as the basis for optimal filter bank design. The scalar non-linear gain-plus-additive noise quantization model can be used to represent each vector quantizer in 2-band subband codec. The validity and accuracy and of this analytic model is confirmed by comparing the calculated model quantization errors with actual simulation of the optimum LBG vector quantizer. Numerical design examples for the optimum separable paraunitary filter banks are suggested in this paper.

## 1. Introduction

Filter banks that operate on 2-dimensional signals find application primarily in subband coding of images and video data. In one dimension, downsampling a sequence by a factor of $M$ is accomplished by simply retaining every $M$-th sample and discarding the rest. The 2-dimensional multirate filter bank is characterized by the subsampling matrix $M$. Thus the filter design problem necessarily depends on the choice of the sampling matrix. For the separable filter bank which implies a diagonal matrix $D$, the filter design is a conceptually trivial extension of most of the results from the 1-D filter bank. These can be easily extended to the multiple dimension in a separable fashion along each dimension. However, for the nonseparable case, it is not that simple and the effect of the sampling matrix $D$ must be carefully considered into the design problem. We compute the mean squared reconstruction error which depends on $N$ the number of entries in each codebook, and $k$ the length of each codeword-(that is, the average bit rate), and on the filter coefficients in this suggested paper. This mean squared error is then minimized subject to the perfect reconstruction constraints and the total bits for the entire bank. The algorithm gives optimum filter coefficients, compensator, gains, and the bit allocations. Numerical design examples for the optimum separable paraunitary is developed in this paper. Also, theoretical results are confirmed by extensive Monte Carlo simulation.

## 2. Modeling of Vector Quantizer for 2 dimensional Filter Banks

The system under study is shown in Fig. 1. The codebook for the vector quantizer for each channel is constructed from the Linde-Buzo-Gray(LBG) algorithm using LENA, PEOPLE, MIT, GIRL, PORTRAIT images of $x(\underline{n})$ passed through a bank of FIR filters designed under Perfect Reconstruction(PR) constraints.

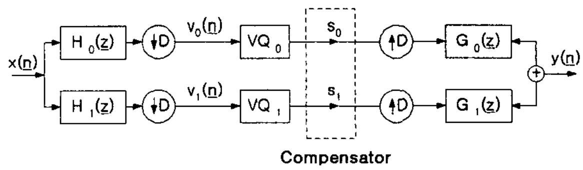



Fig. 1 : 2-dimensional two-channel FB structure using VQ

The overall constraint for the structure is an average $R$ bits/pixel which is to be allocated among these $M$ channels. $\sum_{j=0}^{M-1} R_j = MR$, where $R_j$ is the number of bits allocated to quantizer $Q_j$. Each VQ has a codebook of $N_i$ entries of length $k_i$, and therefore $R_j$ satisfies $R_j = \frac{\log_2 N_j}{k_j}$. In Ref[1], the mean square quantization error for high resolution is given by

$$D^{k}_{VQ}(R) = C(k, m) 2^{-(m/k)R} \left[ \int [p(\underline{v})]^{k/(m+k)} d\underline{v} \right]^{(m+k)/k} \quad (1)$$

where $\underline{v}$ is the signal vector representing $k$ successive values of the input signal $v(\underline{n})$. The vector quantization coefficient $C(k, m)$ is a function of the vector dimension $k$ and of $m$ and represents how well cells can be packed in $k$-dimensional space[2]. The density function $p(\underline{v})$ is the $k$-dimensional joint pdf of the vector process. According to Ref[4], the short-time pdf of a image segment can be approximated by a Gaussian pdf. In Ref[2], the mean squared quantization error averaged over a frame in vector quantization coding can be computed approximately using the asymptotic distortion-rate function derived for a Gaussian random signal,

$$\sigma^2_{\underline{v}} \approx \tau 2^{-2R/k} (\det \Gamma)^{1/k} \quad (2)$$

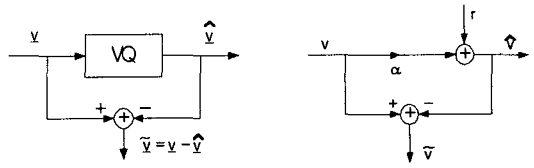
where $k$, $R$ and $\Gamma$ denote respectively the vector dimension, the number of bits allocated to quantizer and the covariance matrix of the input signal; and $\tau$ is a correction factor, $\tau = 2\pi ck(1+\frac{2}{k})^{k/2+1}$ where $c$ is the vector quantization coefficient shown in Ref[2]. Using the Toeplitz distribution theorem[4],

$$\lim_{k \to \infty} \det \Gamma^{1/k} = \exp\left[\left(\frac{1}{2\pi}\right)^2 \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \log_e S_{vv}(e^{j\omega_1}, e^{j\omega_2}) d\omega_1 d\omega_2\right] = \sigma^2_{e,\min}$$

where $S_{vv}(e^{j\omega_1}, e^{j\omega_2})$ is the power spectral density of the random signal $\{v(\underline{n})\}$ and $\sigma_{e,\min}{}^2$ the energy of the minimum prediction error. When the vector dimension $k$ and the predictor order are reasonably large, the quantization error in Ref[2] can be further simplified as ; $\sigma^2_{\underline{v}} \approx \tau 2^{-2R/k} \sigma^2_e$ where $\sigma^2_e$ is the variance of the prediction error sequence using a finite memory optimal predictor. In our case, $R$ is small and $k=9$ is also small. Therefore we introduce an empirically determined correction factor $\delta$ which depends on $R$ and $k$. In Ref[2], introducing correction factor $\delta$ complicates the writing of optimum bit allocation and we know that correction $\delta$ factor gives a little influence on system performance from simulation. So we ignore $\delta$ value. Next, $\sigma^2_e = E\{(v - \hat{v})^2\}$ is expressed an $\sigma^2_e = \gamma^2_v \sigma^2_v$, where $\sigma^2_v$ is energy in the signal input to the quantizer and $\gamma^2_v$ is the 2-dimensional "spectral flatness measure" [4] of $v$. The VQ mean squared error is now

$$\sigma^2_{\tilde{\underline{v}}} = \tau 2^{-2(R/k)} \gamma^2_v \sigma^2_v \quad (3)$$

The pdf optimized scalar quantizer shown in Fig. 2 can be used to represent the optimized VQ[3].




(a) (b)

Fig. 2 : (a) Vector quantizer.
(b) Gain-plus-additive noise model.

## 2.1. 2-D Optimum Paraunitary Filter Bank

The paraunitary filter bank satisfies perfect reconstruction(PR) by choosing the synthesis polyphase matrix

$$G_p(\underline{z}) = \underline{z}^{-\mu} H_p(\underline{z}^{-1}) \quad (4)$$

from the sufficient PR condition

$$P(\underline{z}) = G_p^T(\underline{z}) H_p(\underline{z}) = \underline{z}^{-\mu} I. \quad (5)$$

For this case, the analysis/synthesis polyphase matrices are lossless such that

$$\tilde{H}_p(\underline{z}) H_p(\underline{z}) = I \text{ and } \tilde{G}_p(\underline{z}) G_p(\underline{z}) = I \quad (6)$$

where $\tilde{H}_p(\underline{z}) \equiv H_p^T(\underline{z}^{-1})$. From Ref[3] we have the spatial-domain perfect reconstruction(PR) condition

$$\rho_{rs}(D\underline{n}) = \langle h_r(\underline{k}), h_s(D\underline{n} + \underline{k}) \rangle = \delta_{rs} \delta(\underline{n}). \quad (7)$$

Eq. () implies the following orthogonality conditions

$$\sum_{\underline{k} \in \Lambda} h_r(\underline{k}) h_r(D\underline{n} + \underline{k}) = \delta(\underline{n}), \quad \sum_{\underline{k} \in \Lambda} h_r(\underline{k}) h_s(D\underline{n} + \underline{k}) = 0 \text{ for } r \neq s. \quad (8)$$

Eq.(8) tells us that each filter $\{h_r(\underline{k})\}$ is orthogonal to its translates with respect to the lattice and $\{h_r(\underline{k})\}$ is orthogonal to $\{h_s(\underline{k})\}$ and to all translates of $\{h_s(\underline{k})\}$ in the lattice[3]. This is a lattice extension of paraunitary perfect reconstruction(PR) condition. Note that Eq.(7) asserts that the PR condition must be satisfied for all points $\underline{m} \in D\underline{n}$ of correlation function while in 1-D, the condition is defined for every $M$th sample of the correlation function. For 2-D two-channel paraunitary filter bank with $H_0(z_1, z_2)$ of size $L_1 \times L_2$ FIR filter( $L_1 \times L_2 =$ even), we choose

$$H_1(z_1, z_2) = z_1^{-(L_1-1)} z_2^{-(L_2-1)} H_0(-z_1^{-1}, -z_2^{-1}) \Leftrightarrow$$

$$h_1(n_1, n_2) = (-1)^{(n_1+n_2+1)} h_0(L_1 - 1 - n_1, L_2 - 1 - n_2)$$

to satisfy the paraunitary PR conditions. Then the synthesis filters are spatially-reversed versions of the analysis filters

$$G_0(z_1, z_2) = z_1^{-(L_1-1)} z_2^{-(L_2-1)} H_0(z_1^{-1}, z_2^{-1}) \leftrightarrow$$

$$g_0(n_1, n_2) = h_0(L_1 - 1 - n_1, L_2 - 1 - n_2)$$

$$G_1(z_1, z_2) = z_1^{-(L_1-1)} z_2^{-(L_2-1)} H_1(z_1^{-1}, z_2^{-1}) \leftrightarrow$$

$$g_1(n_1, n_2) = h_1(L_1 - 1 - n_1, L_2 - 1 - n_2) = (-1)^{(n_1+n_2)} h_0(n_1, n_2).$$

Above terms imply that the synthesis polyphase matrix is also lossless and we can rewrite the paraunitary PR condition in Ref[3] in terms of the synthesis filters. Then the orthogonality conditions remove the cross-correlation effects between the channels in the expression for the total output MSE in Ref[3]. Thus the 2-D version of the MS quantization error formula reduces to

$$\sigma_{y_e}^2 = \sigma_d^2 + \sigma_n^2$$

$$\sigma_d^2 = \frac{1}{M} \sum_{i=0}^{M-1} (a_i s_i - 1)^2 \sigma_{v_i}^2, \quad \sigma_n^2 = \frac{1}{M} \sum_{i=0}^{M-1} s_i^2 \sigma_{r_i}^2 \tag{9}$$

where $\sigma_{v_i}^2 = \sum_{\underline{k}} \sum_{\underline{l}} h_i(\underline{k}) h_i(\underline{l}) R_{xx}(\underline{k} - \underline{l})$.

For the optimal compensator $\{s_i^*\}$, we minimize Eq.(9) by setting $\frac{\partial \sigma_{y_e}^2}{\partial s_i} = 0$, and we can again show that $s_i^* = 1$. Again the paraunitary structure with no compensation is optimum.

## 3. Design Examples and Experimental Results

Design examples for the optimal paraunitary filter bank are presented for 2-D filter bank. For the separable FB case, we have a four channel tree like structure so that each row and column operation can be applied to the horizontal and vertical directions respectively. For this case, we design the analysis/synthesis FB of equal size 6×6 where each directional 1-D prototype filter is of length 6. Five images, LENA, PEOPLE, GIRL, PORTRAIT, and MIT are used for training image for VQ. Thus we will show the analysis and simulation mainly on the 256×256 Lena image as a example. The optimization algorithm is again based on the exhaustive search of all possible bit allocations constrained by the total number of bits. For each bit combinations, we compute the optimal filter coefficients, optimal compensators, and the associated MSE. We choose the one with the smallest MSE among them. We assume a globally stationary source model for the test images which is a rather strong assumption, but it makes the analysis simpler. Then the separable autocorrelation function of AR(1) source can be written as $R_{xx}(m,n) = \sigma_x^2 \rho_h^{|m|} \rho_v^{|n|}$ where the input variance $\sigma_x^2 = \frac{1}{(1-\rho_h^2)(1-\rho_v^2)}$, and $\rho_h$, $\rho_v$ are directional correlation coefficients. Note that the first-order correlation coefficients $\rho_h$, $\rho_v$ can be easily determined from $\rho_h = R_{xx}(1,0)$, $\rho_v = R_{xx}(0,1)$ with respect to normalized input variance $\sigma_x^2$=1.

### 3.1. Calculation Procedures for Optimal Filter

1.) We allocate bits to VQ of low-pass filter and high-pass filter. For separable filter, we allocate bits 4-band filter after downsampling and we allocate bits 2-band filter after downsampling for nonseparable filter. We fix bit rate $R$, vector dimension $k$, input image correlation coefficient $\rho_h$, $\rho_v$. Calculate $\tau = 2\pi c k (1+\frac{2}{k})^{2/k+1}$ from Tables[2]. Set $V(0)$ to large number.

2.) Choose $h_0(n_1, n_2)$ satisfying Perfect Reconstruction(PR) paraunitary filter for optimal paraunitary filter.

3.) Calculate

$$S_{xx}(e^{j\omega_1}, e^{j\omega_2}) = \frac{(1-\rho_h^2)(1-\rho_v^2)}{(1+\rho_h^2 - 2\rho_h \cos(\omega_1))(1+\rho_v^2 - 2\rho_v \cos(\omega_2))} \sigma_x^2$$

Calculate $S_{vv}(e^{j\omega_1}, e^{j\omega_2}) = |H(e^{j\omega_1}, e^{j\omega_2})|^2 S_{xx}(e^{j\omega_1}, e^{j\omega_2})$.

Calculate spectral flatness

$$\gamma_v^2 = \frac{\exp[(\frac{1}{2\pi})^2 \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \log_e S_{vv}(e^{j\omega_1}, e^{j\omega_2}) d\omega_1 d\omega_2]}{\frac{1}{2\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} S_{vv}(e^{j\omega_1}, e^{j\omega_2}) d\omega_1 d\omega_2}.$$

Calculate $\alpha = 1 - \tau 2^{-2(R/k)} \gamma_v^2$.

Calculate $\sigma_r^2 = \alpha(1-\alpha)\sigma_v^2$.

4.) The MSE of paraunitary filter banks is shown in Eq.(9)[3].

5.) The filter solutions was obtained by using IMSL Library(DNCONF) software. This package solves a general nonlinear constrained mimimization problem using the successive quadratic programming algorithm and a finite difference gradient. Thus we obtain optimum filter $h_0(n_1, n_2)$ and MSE.

6.) Is $(MSE)^i \le (MSE)^{i-1}$ ? If yes, go to step 3 and if not, stop.

7.) Finally, we obtain optimal filter coefficient and minimum MSE.

### 3.2. Separable Filter Bank Case

Table 1 shows the optimum bits allocated to each channels in bits/pixel, optimum compensators, the simulated output MSE for the separable paraunitary for LENA image. For the paraunitary case, the optimum compensator was unity, so it is not shown on the table. From the tables, we know the output MSE of both filter banks is almost the same at these average bit rates. Table 2 shows the optimum directional 1-D paraunitary LPF $h_0(n)$( HPF,$h_1(n) = h_0(5-n)$) for LENA image. We note that the optimum separable analysis and synthesis FB can be easily implemented with the directional 1-D filters as follows,

$h_1(n) = (-1)^{n+1} h_0(5-n)$

$g_0(n) = h_0(5-n)$

$g_1(n) = (-1)^n h_0(n)$

The horizontal higher frequency band *LH* has horizontal edges and vertical higher frequency bands *HL* has vertical edges and the highest frequency band *HH* has the diagonal edge information.

1.) Input image is filtered with row 1-D analysis filter and downsampled by 2 in row direction.

2.) Filtered with column 1-D analysis filter and downsampled by 2 in column direction.

3.) *Allocated bits to each 4 band($R_0$, $R_1$, $R_2$, $R_3$) output* channels.

4.) Coded and decoded using VQ codebook.

4.1.) We use the LBG algorithm to create the codebook.

4.2.) Fix vector dimension $k$=9(3 by 3 pixels), bit rate $R$.

4.3.) Training images, LENA, MIT, GIRL, PEOPLE, PORTRAIT are used for codebook design.

5.) Compensated by compensation gain $s_i$.

6.) Upsamplered by and filtered with 1-D synthesis filter and summed in column direction.

7.) Upsampled and filtered with 1-D synthesis filter and summed in row direction. Thus, we obtained the reconstructed image.

## 4. Conclusions

As seen from the Table 2, the paraunitary FB is less sensitive to variation of the average bit rate $R$ bits/pixel. Fig. 3 shows the two images: (a),(b) are reconstructed output and the error image for the separable paraunitary filter at the average bit rate $R$ =0.36 bits/pixel. We see that the quantization effects are most visible near the LENA's hat and the edges. We have developed the optimum paraunitary FB for separable cases and demonstrated these optimal filters on the actual test images. The simulation results on LENA image show that the quantization is most significant near the edges. This clearly calls for the separate processing of the edges at the transmitter side. A possible scheme might be to attach an edge detector at the high frequency channel before the analysis filter. For this case, we will, of course, lose the perfect reconstruction property of the filter bank even in the absence of quantizers and compensators. We can apply this methods to both nonseparable and biorthogonal filter banks.

## 5. References


[1] A. Gersho, ''Asymptotocally optimal block quantization,'' in IEEE Trans. on Information Theory, vol. IT-25, No. 4, pp. 373-380, July 1979.

[2] V. Cuperman, ''On adaptive vector transform quantization for speech signal,'' in IEEE Trans. on Comm., vol. 37, No. 3, pp. 261-267, April 1989.

[3] I. Jee and R. A. Haddad, ''Optimum design of vector-quantized subband codecs,'' in IEEE Trans. on Signal Processing, vol. 46, No. 8, pp. 2239-2243, Aug. 1988.

[4] N. S. Jayant and P. Noll, '' *Digital Coding of* Waveforms, Englewood Cliffs, NJ, Prentice-Hall, 1984.

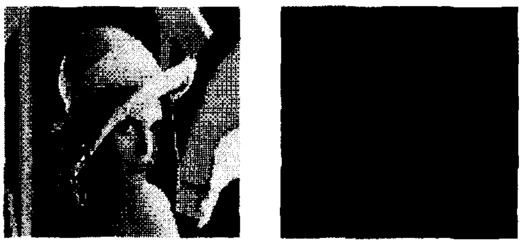
[5] G. Karlsson and M. Vetterli, 'Theory of two-dimensional multirate filter banks,'' IEEE Trans. Acoustics and Speech and Signal Processing, vol. 38, No. 6, pp. 925-937, June 1990.


| R | $R_0$ | $R_1$ | $R_2$ | $R_3$ | MSE |
|---|---|---|---|---|---|
| 0.25 | 6 | 1 | 1 | 1 | 0.08882 |
| 0.31 | 8 | 1 | 1 | 1 | 0.06200 |
| 0.36 | 10 | 1 | 1 | 1 | 0.05206 |

Table 1 : Optimum bit allocations and simulated MSE for separable paraunitary FB

| R | $ph_0(0)$ | $ph_0(1)$ | $ph_0(2)$ | $ph_0(3)$ | $ph_0(4)$ | $ph_0(5)$ |
|---|---|---|---|---|---|---|
| 0.25 | 0.33088479 | 0.80630637 | 0.46225321 | -0.134965 | -0.085251 | 0.0349847 |
| 0.31 | 0.33088570 | 0.80630600 | 0.46225299 | -0.134964 | -0.085253 | 0.0349855 |
| 0.36 | 0.33088570 | 0.80630600 | 0.46225299 | -0.134964 | -0.085251 | 0.0349861 |

Table 2 : 1-D directional LPF $ph_0(n)$ for optimum separable paraunitary FB for LENA image



(a) (b)

Fig. 3 : (a) Reconstructed output image, (b) error image for separable parauniary FB at R=0.36 bit/pixel, MSE=0.05206