Finite State Lattice Vector Quantization
for Wavelet-Based Image Coding

J.Q. Ni
jqni@hkueee.hku.hk

K.L. Ho
klho@hkueee.hku.hk

K.W.Tse
kwtse@hkueee.hku.hk

Department of Electrical & Electronic Engineering
The University of Hong Kong
Pokfulam Road, Hong Kong

Abstract - It is well known that there exists strong energy correlation between various subbands of a real-world image. A new powerful technique of Finite State Vector Quantization (FSVQ) has been introduced to fully exploit the self-similarity of the image in wavelet domain across different scales. Lattices in R^N have considerable structure, and hence, Lattice VQ offers the promise of design simplicity and reduced complexity encoding. The combination of FSVQ and LVQ gives rise to the so-called FSLVQ, which is proved to be successful in exploiting the energy correlation across scales and simple enough in implementation.

I. Introduction

Recently wavelet based image coding has become the domain of extensive research. Generally a K-level wavelet decomposition of a nature image yields 3K+1 subbands, forming three K+1 level directional pyramids. A simple observation of the high frequency subbands (HFS) reveals the fact that the bulk of the energy in HFS's is concentrated more-or-less in the vicinity of areas which reflect the edge activity of the original image. This implies that the energy distribution in HFS is not uniform and those areas of HFS's that contain most of the energy must be encoded more finely than the rest. It is also well known that there exist strong energy correlations between each subband images with the same orientation, which appears as self-similarity of the image in wavelet domain across different scales. Therefore it is possible to predict the activity of one band using the other band in the same orientation. The above two observations motivated the principle of Finite State Quantization in wavelet domain, which first proposed by T.Naveen and J.Woods in their original work[5], where they apply the powerful technique finite state scalar quantization to the multiscale subband coding.

It has been shown that vector quantization can offer better coding performance compared to scalar quantization, mainly due to its inherent advantage of fractional bit allocation among the vector components. Lattice vector quantization (LVQ) is a type of VQ where the codebook is designed based on regular lattices. Lattice VQ offers the promises of design simplicity and reduced complexity encoding. The goal of the paper is to extend the work in [5] by combining the FSVQ with LVQ, which

gives rise to the so-called Finite State Lattice Vector Quantization (FSLVQ) and provides further exploration of the energy correlation between spatial subbands of image while preserving simplicity in implementation.

II Finite State Vector Quantization

II.1 *The Fundamental of FSVQ*

An FSVQ can be viewed as a collection of K separate vector quantizers together with a selection rule that determine which of the K VQ codebooks is used to encode the current input vector into a channel index. The key issue involved in FSVQ is the design of next-state function. It can be tackled by applying a classifier function to the reproduction of input vector, i.e., find $s_{n+1} = V(\hat{x}_n)$, which is tractable for the decoder. In summary, we can keep the quantizer designed for the the idealized or omniscient state, which is the so-called omniscient design of FSVQ.

II-2 *The FSVQ Configuation in Wavelet Domain*

As we state in the previous section, in wavelet transform domain the activity of one subband can be predicted by another subband of similar orientation. The element at the coarse scale is called the parent, and all elements corresponding to the same spatial location at the next finer scale of similar orientation are called children (See Fig.1). With the exception of lowest frequency subband, all parent have four children. While for the lowest frequency subband, each parient has only three children. The configuation is idealized for FSVQ approach, i.e., a spatial DPCM for the lowest frequency subband and finite state VQ for the higher frequency subbands. The non-uniform distribution of energy in a subband is reflected by the state of vector quantizer. We implement the FSVQ with a top-to-bottom fashion in the wavelet pyramid, i.e., the lower frequency subbands are quantized first. A special function V is designed to classify the quantizer outputs, while the class of the output element serves as the state of its children. It is worth pointing out that the state determination at both encoder and decoder is based on the decoded elements. Thus the decoder can track the next-state rule of encoder and the quantization process has memory, which constitutes the FSVQ.

III Lattice Vector Quantization

III.1 *The principle of Lattice Vector Quantization*

The considerable structure associated with a regular lattice and the fast quantization algorithm that this structure implies, have been stimulating for the use of lattice as vector quantizer[1]. This quantization is more attractive than conventional nonuniform vector quantization in that when followed by a variable length entropy encoder it provides performance comparable to nonuniform vector quantizer with a much reduced computational complexity. A lattice is defined as a set of vectors

$$\Lambda = \{ \mathbf{x} : \mathbf{x} = u_1\mathbf{a_1} + u_2\mathbf{a_2} + \cdots + u_N\mathbf{a_N} \} \qquad (\text{III-1})$$

where $\mathbf{a}_i$, $i = 1, 2, \cdots, N$ are the basis vectors of the lattice and the u_i are integers. The lattice VQ is constituted by selecting some of the lattice points $\mathbf{x}$ to be the output points $\mathbf{Y_i}$ and forming Voronoi regions Ω_i associated with these output points so that if a source vector $\mathbf{X} \in \Omega_i$ then $\mathbf{Y_i} = \mathbf{Q}(\mathbf{X})$.

III.2 *Pyramid Lattice Vector Quantization*

It is well known that the Laplacian model is well fitted to the pdf of subband image data. For an L – Dimensional vector $\mathbf{x} = (x_1, x_2, \cdots, x_L)^\tau$ consisting of i.i.d. Laplacian variables x_i with variance $2 / \lambda_i^2$ and mean zero, the joint pdf of $\mathbf{x}$ is

$$f(\mathbf{x}) = (\prod_{i=1}^{L} \frac{\lambda_i}{2}) \exp\{-\sum_{i=1}^{L} \lambda_i |x_i|\} \qquad (\text{III-2})$$

The contours of constant density satisfy $\sum_{i=1}^{L} \lambda_i |x_i| = constant$. We then define a "weighted pyramid" as :

$$S(L, \lambda, C) = \{\mathbf{x} : \frac{1}{L}\sum_{i=1}^{L} \lambda_i |x_i| = C\} \qquad (\text{III-3})$$

The asymptotic equipartition principle of information theory indicates that for large encoding rate and uniformly distributed codewords, all codewords close to $S(L, \lambda, C)$ are roughly equally likely [2]. This motivates us to construct a VQ/noiseless code by using a lattice to define the uniformly distributed codewords in R^L and a suitable enumeration code to assign equal-length binary codewords to each lattice points on the surface of same pyramid $S(L, \lambda, C)$. Usually suitable scaling and rounding are introduced to make the enumeration code feasible in practice.

Since the vector components in our project are from the same class in a subband and can be viewed to have the same variance, a simplified pyramid can be defined as follow:

$$S(L, r) = \{\mathbf{x} \in R^L : \sum_{i=1}^{L} |x_i| = r\} \qquad (\text{III-4})$$

where the pyramid "radius" r is the l_1 norm of $\mathbf{x}$. We focus on cubic lattice Z^N in our project for simplicity,

although other lattice can also be investigated. Let the integer function $f(x)$ to be the closest interger to a real x. The number of cubic lattice points on $S(L, r)$ is $N(L, r)$. The introduced scaling factor or "quantization step" Δ is used to adjust the trade-off between entropy and distortion. Then the encoding process with LVQ for an input L – Dimensional vector $\mathbf{x}$ is as follows:

1) LVQ with cubic lattice:

The input vector $\mathbf{x}$ is scaled with Δ, and quantized as $\tilde{\mathbf{y}} = f(\mathbf{x} / \Delta)$.

2) Enumeration

   a) The l_1 – norm of $\tilde{\mathbf{y}}$, $r = \|\tilde{\mathbf{y}}\|_1$ is computed.

   b) r is encoded with entropy code (variable length) C as $c'(r)$.

   c) The enumeration code $C''(r)$ developed in [4] is used to encode $\tilde{\mathbf{y}}$ with a binary codeword c' of length $\lceil \log_2 N(L, r) \rceil$ b

   d) The union of the two sub-codeword $c = (c', c'')$ is then stored and transmitted.

IV The FSLVQ Algorithm

IV-1 *The Design of FSLVQ for Wavelet-Based Image Coding*

A J-level wavelet decomposition of image yields a J-layer wavelet pyramid structure $\{P_J, (D_j^1)_{1 \le j \le J}, (D_j^2)_{1 \le j \le J}, (D_{1 \le j \le J}^3)\}$, where P_J, D_j^k denotes the LFS and detail in subband (j, k), respectively.

We classify the elements in a subband into S classes based on its local variance. Although the nonuniform classification gives superior performance, we adopt a "equipopulous classes" approach[7] in our project for the interest of simplicity, which works well, provided that the number of class is sufficient large($S = 4$ in our project). This classification then determines the state of the quantizer for a higher frequency subband. The process of classification for the FSLVQ design is described as follows:

> **Procedure Classification(P, J ,S)**
> **begin**
> **decompose the input image P with WT; implement the DPCM on P_J, and obtain the open-loop prediction error.**
> **for level j=J to 1**
> **for pyramid direction k=1 to 3**
>   **break the parent subband into block of size $m \times m$, and classify the blocks into S equipopulous classes based on their loacl variance; partition the blocks in subband (j,k) to obtain the**
> $D_{j,s}^k = \{x \in D_j^k, Class(p(x)) = s\}$;
> **next**
> **next**
> **end**

Note: x and $p(x)$ are the block in D_j^k and the parent of it, respectively.

As in all subband coding schemes, one must determine how to allocate bits among the subbands. This issue can be formulated as the following constrained optimization problem:

$$\text{Min}\left\{ \varepsilon_J(R_J) + \sum_{j=1}^{J}\sum_{k=1}^{3}\frac{1}{S}\sum_{s=1}^{S}\varepsilon_{j,s}^k(R_{j,s}^k) \right\} \qquad (\text{IV-1})$$

subject to bit rate constraint

$$(\frac{1}{4^J}R_J + \sum_{j=1}^{J}\frac{1}{4^j}\sum_{k=1}^{3}\frac{1}{S}\sum_{s=1}^{S}R_{j,s}^k) \leq R_{T\arg et} \qquad (\text{IV-2})$$

where $\varepsilon_{j,s}^k$ and $R_{j,s}^k$ denote the distortion and rate of s-class in subband (j,k).

The conventional approach is to obtain the optimum bit allocation by a Lagrange multiplier method [5]. However, the results are not constraint to non-negative bit allocation. A novel adaptive scheme developed by authors[6] can be extended to design the FSLVQ quantizer. Since the rate and distortion trade-off of lattice vector quantizer is adjusted by the scaling factor Δ, the overall distortion in (IV-1) can be expressed as :

$$\varepsilon = D(\Delta_J, \Delta_{j,s}^k) \qquad (\text{IV-3})$$

Then the aforementioned optimization problem is equivalent to finding a set of optimal scaling factor $\left\{ \Delta_J, \left(\Delta_{j,s}^k\right)_{1\leq j\leq J, 1\leq k\leq 3, 1\leq s\leq S} \right\}$ to minimize (IV-1) subject to (IV-2).

The basic concept of the scheme is described as follows:starting from a set of initial large scaling factor $\left\{ \Delta_J, \left(\Delta_{j,s}^k\right)_{1\leq j\leq J, 1\leq k\leq 3, 1\leq s\leq S} \right\}$ for every class in each subband, corresponding to low bit rate and high restruction distortion, the algorithm decreases one entry of the set of scaling factor at a time until a target bit rate is reached. The quantity $\dfrac{-\Delta\varepsilon}{\Delta R}$, i.e., the ratio between the decrease in reconstruction error and the increase in bit rates, is used as a criterion to reflect the rate-distortion trade-off. Then the procedure of the quantizer design can be described as:

$$\max_{j,k,s} \max_{\Delta_{j,s}^k} \frac{-\Delta\varepsilon}{\Delta R}\bigg|_{\Delta_{j,s}^k \rightarrow \overline{\Delta}_{j,s}^k} \qquad (\text{IV-4})$$

The detailed design procedure can be found in [6]. When performing the design algorithm, one can estimate the reconstruction error through the quantization error Via (IV-1). There also exists a convenient way for fast rate estimation. We notice that for a given weight vector W (unit weight in our project), the length of the enumeration codeword can be precomputed for each l_1 norm and stored as a table. Thus simply computing the l_1 norm for each vector to be encoded allows a rapid evaluation of the number of bits required to encode a string of source vectors. If there are M blocks belonging to state s in

subband (j,k), the estimated average bit rate for a particular scaling factor $\Delta_{j,s}^k$ can be expressed as:

$$\Re_{j,s}^k = \frac{1}{4^J SL}\left[H(r_m) + \frac{1}{M}\sum_{m=1}^{M}\lceil \log_2(N(L,r_m)) \rceil \right] \qquad (\text{IV-5})$$

where $H(r_m)$ is the entropy of l_1 norm r_m.

IV-2 *The Implementation of FSLVQ*

The FSLVQ quantizer designed through the process outlined above can be used to encode the image in wavelet domain. The details of encoding and decoding are given in the following procedures

procedure ENCODER(P, $R_{t\arg et}$,J,S)
begin
 decompose the input image P into J-lvel with WT;
 enocde the LFS subband with DPCM; divide the quantized prediction errors into block of size $m \times m$ and classify the blocks into S equipopulous class;
  for level j=J to 1
  for pyramid direction k=1 to 3
  for state s =1 to S
  break the subband (j,k) into blocks of size $m \times m$; determine the set of blocks for state s based on their parents and encode them with LVQ designed for (j,k,s).
  next
  classify the quantized blocks in subband (j,k) into S equipopulous classes.
 next
 next
end

Similarly we have the procedure DECODER.

Computer simulation was carried out to show the feasibility of the proposed FSLVQ algorithm for wavelet-based image coding. We implement the FSLVQ algorithm with 4-class on the standard monochrome Lena image of size $512 \times 512 \times 8$. In our simulation we break the subband image into nonoverlapping blocks of size 4×4 and constitute the 16-D vectors for LVQ. The classification of subband image is also based on the local variance of the 4×4 blocks. The overhead information that needs to send to the decoder includes the scaling factor for LVQ of each class in each subband. Our simulation results are compared to Naveen's FSSQ coding scheme[5] in terms of R-D performance(see fig. 2). The improvement in PSNR with the proposed FSLVQ (4-calss) over the FSSQ (4-class) is nearly 0.2db. One can improve the R-D performance by encoding the LFS subband with DCT ECLVQ and predicting the classes in next higher frequency subband based on the quantized ac-coefficients. Further improvement in performance can be expected if we adopt large block size or other more compact lattice such as E_8 and Λ_{16}.

V. Conclusion

In this paper, we have proposed a Finite State Lattice Vector Quantization scheme for wavelet-based image coding. Finite-State VQ have been proved to be

successful in exploiting the self-similarity of the image in wavelet domain across different scale. Lattice VQ offers the promise of design simplicity and reduced complexity encoding. The proposed FSLVQ combines both the efficiency of FSVQ in exploiting energy correlation and simplicity of LVQ in implementation. A novel adaptive scheme developed by authors[6] is also used to design the FSLVQ quantizer, which adapts to the actual statistical of input images. Finally, comparisons with the earlier version of FSVQ, i.e., Naveen's suband-FSSQ[5], which is very efficient in exploiting the energy correlation among subbands with scalar quantizer, favour the proposed FSLVQ quantizer.

Reference

[1] M. Barlaud , M. Antonini , " Pyramid Lattice Vector Quantization for Multiscale Image Coding," IEEE Trans. on Image Processing, Vol. 3, No. 4, pp.367-380, July 1994.

[2] T.R. Fischer, " A Pyramid Vector Quantizer," IEEE Trans. on IT, Vol. IT-32, pp 568--583, 1986.

[3] T.R. Fischer, " Entropy-Constrained Geometric Vector Quantization for Transform Image Coding," in Proc. IEEE ICASSP-91, pp. 2269-2272.

[4] T.R. Fischer and J. Pan " Enumeration Encoding and Decoding Algorithms for Pyramid Cubic Lattice and Trellis codes," IEEE Trans. on IT, Vol.41, No.6, pp.2056-2061, 1995.

[5] T. Naveen and J. Woods, " Subband Finite State Scalar Quantization, " IEEE Trans On Image Processing, Vol. 5, No. 1, pp. 150-155, 1996.

[6] J.Q. NI, K.L. Ho and K.W. Tse, " Rate-Constrained Adaptive Quantization Scheme for Wavelet Pyramid Image Coding," In Proc. ISCAS-96, pp.636-639, 1996.

[7] W.H. Chen, " Adaptive Coding of Monochrome and Color Images," IEEE Trans. on Com., Vol.25, pp. 1285-1292, Nov. 1977.
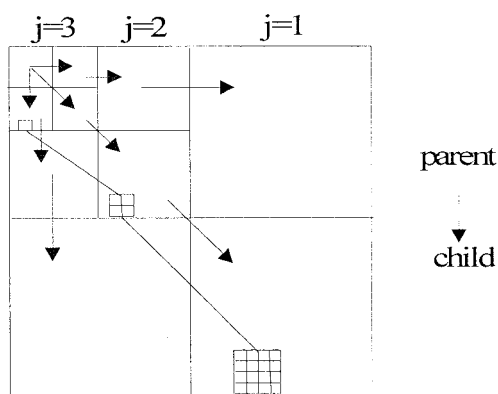
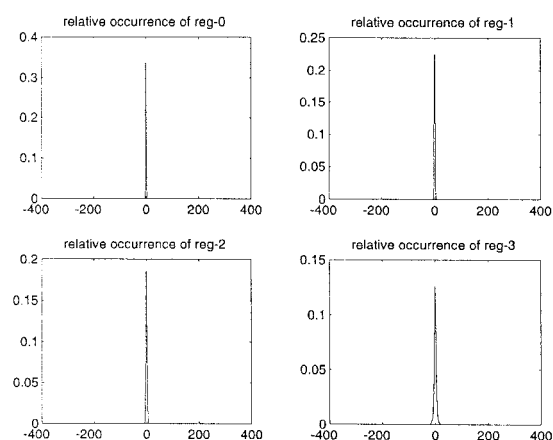Fig.1 3-level wavelet dexomposition and parent-chid relation

Fig.2 Relative occurrence of 4-class for one subband;reg-3 has the largest variance
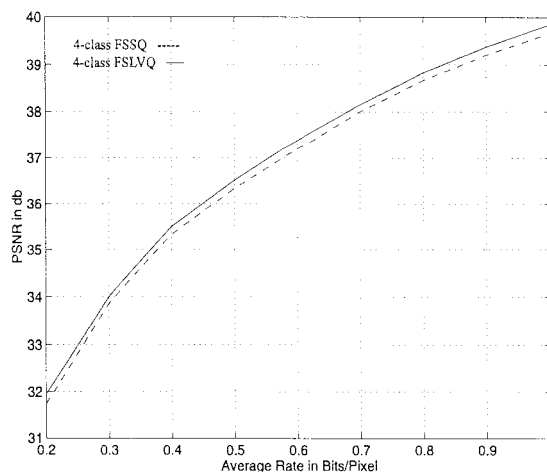

Fig.3 Performance of 4-class FSSQ and 4-class FSLVQ for monochrome 512*512 Lena