JZUS

A high fidelity VQ coding algorithm with
region adaptive subbanding

YANG Zhong-qi (杨仲琦), BAI Zhi-can (白直灿), LI Dong-mei (李冬梅), YANG Chang-sheng (杨长生)[†]

(*Department of Computer, Zhejiang University, Hangzhou 310027, China*)

[†]E-mail: csyang@hzcnc.com

Abstract: Image subbands can be obtained by using filterbank. Traditional compression method uses direct entropy coding for each subband. After studying the energy distribution in image subbands, we proposed a vector quantization (VQ) coding algorithm to image subband. In the algorithm, vector quantizers were adaptively designed for high-frequency bands in an image. In particular, the edges of the image were examined and fewer bits were assigned to high-energy regions. The experimental result showed that the algorithm had higher SNR and higher compression ratio than possible by traditional subband coding, JPEG and JPEG 2000.

Key words: Image compression, Subband, Vector quantization (VQ)
doi:10.1631/jzus.2005.A0043   **Document code:** A   **CLC number:** TN919.81

INTRODUCTION

Digital image compression technique plays a vital role in communication, media storage, and data issue fields, and is one of the most active fields of information techniques.

This paper proposes a near-lossless compression algorithm which applies a high fidelity vector quantization method to image subbands. The original image was first separated by filterbank into seven subbands. The regions around the edges in each subband were then extracted, and divided by their orientations into several sub-regions. In the third step, different codebook (vector quantizer) was designed for different region, and fewer bits were assigned to regions of higher-energy in the subband. Better compression ratio and fidelity were achieved.

ENERGY DISTRIBUTION IN IMAGE SUBBANDS

The properties of the lowest frequency subband (LFS) are similar to those of the original image, but all the highest frequency subband (HFS) images have small correlation coefficients between pixels within a subband and negligible cross-correlation coefficients between subbands. The former is described in Tanabe and Farvardin (1992). The energy distribution of the HFS is described briefly in this section. The reader can be referred to Kwon and Chellappa (1998) for details.

A traditional image compression scheme based on a filterbank splits the signal into M subbands as shown in Fig.1. The input signal defined by $x(n) = \begin{cases} 1, & n = 0,1,2,\cdots \\ 0, & \text{otherwise} \end{cases}$, has the following anti-symmetric property.

Property 1   Let the input signal, $x(n)$, be a high-pass decimator or a band-pass decimator, with the decimated signal, $x_i(n)$, being antisymmetric for $n \neq 0$, that is, $x_i(n) = -x_i(-n)$, $n>0$.

Corollary   Given a $(2N+1)$-length one-dimensional discrete-time signal, $x(n)$, $-N \leq n \leq N$. If the input signal, $x(n)$, is a high-pass decimator or a band-pass decimator, the mean of $x_i(n)$ is equal to $\dfrac{x_i(0)}{2N+1}$, where

$-N \leq n \leq N$.

Let $k_i = M$ and $H_i(\omega) = \begin{cases} 1, & \dfrac{i-1}{M}\pi \leq |\omega| \leq \dfrac{i}{M}\pi \\ 0, & \text{otherwise} \end{cases}$

for $i=1, 2, \ldots, M$ in Fig.1 with subband tree structure forming an M-band subband tree decomposition. For the M-band subband tree decomposition, we have the following energy packing property.

Property 2   In the M-band subband tree decomposition, the energy of the input signal in the ith subband, $E_{x_i}(M)$, is a decreasing function of i for any M.

The energy ratio of the ith subband is defined by $\eta_{x_i}(M) \equiv \dfrac{E_{x_i}(M)}{\sum_{i=2}^{M} E_{x_i}(M)}$, which represents the ratio of the energy of the ith subband to the total HFS energy.

Corollary $E_{x_i}(M)$, or equivalently $\eta_{x_i}(M)$, decreases monotonically as i increases for any M. It also can be proved that $\eta_{x_i}(M)$ is a monotonically decreasing function of M for fixed i and that more than 90% of the total HFS energy is concentrated in the lower half of the HFS when $M \geq 16$.

Let the compact energy of the ith subband be defined by $\xi_{x_i} \equiv |x_i(0)|^2 + |x_i(1)|^2 + |x_i(-1)|^2$ and the compact energy ratio of the ith subband be defined by $\gamma_{x_i} \equiv \dfrac{\xi_{x_i}(M)}{\sum_{i=2}^{M} E_{x_i}(M)}$, which represents the ratio of the compact energy of the ith subband to the total HFS energy. It was found that the properties of γ_{x_i} are very similar to those of η_{x_i}.

Property 3   The compact energy $\xi_{x_i}(M)$, or equivalently $\gamma_{x_i}(M)$, decreases monotonically as i increases for any M. Also $\gamma_{x_i}(M)$ is a monotonically

decreasing function of M if i is fixed, with more than 80% of the total HFS energy being in the form of compact energy for all M.

In the M-band subband tree decomposition, the energy of the ith subband and the energy of the $(i+1)$th subband have the following relationship.

Property 4   In the M-band subband tree decomposition of the input signal, the relationship between the energy of the ith subband, $E_{x_i}(M)$, and the energy of the $(i+1)$th subband, $E_{x_{i+1}}(M)$, can be represented as

$$E_{x_i}(M) = \cos\left(\frac{\pi}{2^{M+1-i}}\right) E_{x_{i+1}}(M) \text{ for } i=2, 3, \ldots, (M-1).$$

The proofs of the above properties and Corollaries are given in Kwon and Chellappa (1998).

Since the property of the compact energy ratio is similar to the property of the energy ratio, we can conclude that the compact energy of each subband is almost equally distributed up to $i=M-1$. It was found that more than 77% of the total HFS energy is in the form of compact energy.

The energy distribution of the image subbands has the following characteristics:

(1) Energy packing property toward the lower frequency subbands: The energy decreases monotonically as i increases, with more than 90% of the total HFS energy being concentrated in the lower half of the HFS.

(2) Energy packing property toward the edges: Most of the total HFS energy is concentrated on the edges. The Lena image has been found to contain 84.9% of the total HFS energy in its edge-regions.

(3) Directionality of the energy distribution: Analysis of the compact energy of the subedge-regions showed that the horizontal, vertical, and off-diagonal edge-region has relatively more energy in
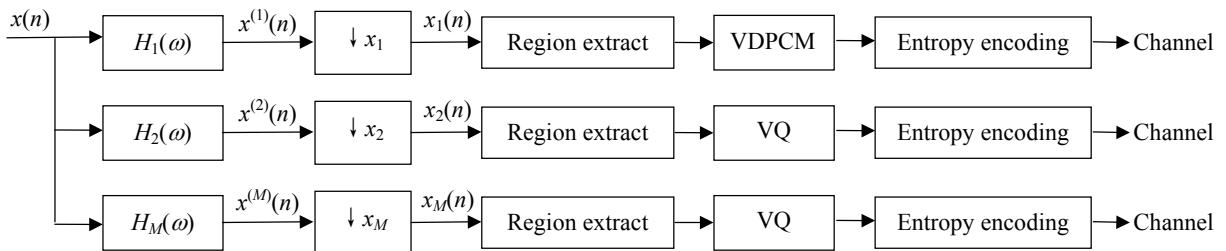

Fig.1 General structure of subband coding system

the horizontal, vertical, and diagonal subbands, respectively, and that the diagonal edge-region has relatively less energy in the diagonal subbands.

VQ CODING ALGORITHM WITH REGION ADAPTIVE SUBBANDING

The above description of the energy distribution leads to the following conclusions.

(1) Based on the energy packing property toward the edges, for efficient transmission of image, the edge-regions should be emphasized;

(2) Based on the directionality of the energy distribution, different bits should be allocated for different sub-regions;

(3) Based on the energy packing property toward the lower frequency subbands, more bits should be assigned to the lower frequency subbands.

Based on the aforementioned conclusions, we present a high fidelity vector quantization (VQ) coding algorithm with region adaptive subbanding. Our encoder is effective at:

(1) Splitting the original image into seven subbands using two-level Walsh filtering;

(2) Detecting the edges for each subband using Canny detector (Canny, 1986);

(3) Partitioning each subband into vectors;

(4) Extracting edge-regions for each subband;

(5) Encoding LFS using VDPCM and encoding HFS using VQ;

(6) Designing entropy coder.

In what follows, we describe in detail each step of our encoder shown in Fig.2.

Subband decomposition

The Walsh-coefficient, $W_x(n)$, can be seen as the frequency domain representation of an N-length one-dimensional discrete-time signal, $x(n)$, which can be decomposed into lower frequency $W_{xL}(n)$ and higher frequency $W_{xH}(n)$ as follows.

$$W_{xL}(n) \equiv \begin{cases} 0, & 0 \leq n \leq \dfrac{N}{2}-1 \\ W_x(n), & \dfrac{N}{2} \leq n \leq N-1 \end{cases}$$

$$W_{xH}(n) \equiv \begin{cases} W_x(n), & 0 \leq n \leq \dfrac{N}{2}-1 \\ 0, & \dfrac{N}{2} \leq n \leq N-1 \end{cases}$$

We can obtain the low-pass signal, $x_L(n)$, and the high-pass signal, $x_H(n)$, through inverse Walsh transform, which has the following property:

Property 5   Two adjacent points of the low-band signal, $x_L(n)$, of a bank of N points are equal with positive sign and two adjacent points of the high-band signal, $x_H(n)$, are equal with negative sign; that is, $x_L(2i)=x_L(2i+1)$, $x_H(2i)=-x_H(2i+1)$ for $i=0, 1, \ldots, N/2-1$.

Based on the above property, the subband can be downsampled losslessly by two. We have the following conclusion:

Property 6   The relation between the low-pass signal, $x_L(n)$, and the high-pass signal, $x_H(n)$, of the signal, $x(n)$, can be represented as $\bar{x}_L(n) + \bar{x}_H(n) = \bar{x}(n)$.

It was proved that the Walsh analysis/synthesis filterbank has ideal reconstruction performance (Manz, 1972; Ching and Goodyear, 1984), so we applied it to realize two-level filtering in our experiment, in which the original image was decomposed into seven subbands using two-level Walsh filtering shown in Fig.3.
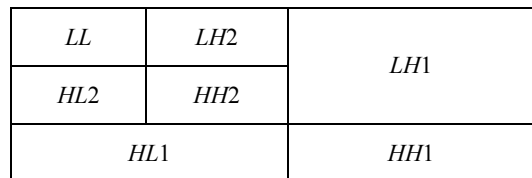
LL	LH2	LH1
HL2	HH2	
HL1		HH1

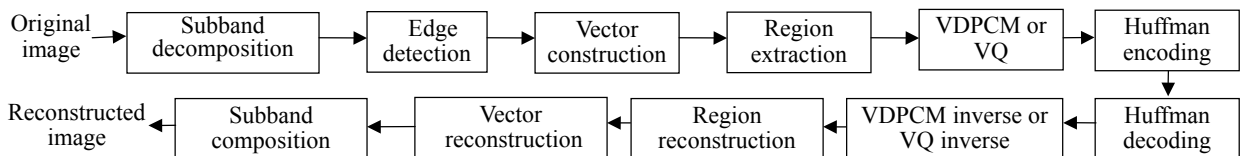**Fig.3 Two-level subband decomposition**

Fig.2 General framework of encoder and decoder

Edge detection and vector-region extraction

The horizontal, vertical, diagonal, and off-diagonal edges for each subband were detected using the Canny detector (Canny, 1986); then the recursive thresholding method (Kwon and Chellappa, 1998) using multiple one-dimensional histograms of several texture features was used to extract tex-tured-regions.

The edge-region extraction algorithm of Kwon and Chellappa (1998) was extended to the following vector-edge-region extraction algorithm.

Let the vector-edge-region of the (i,j)-th sub-band be Λ_{ij}; detect the edges $\{x(m, n)\}$ for the (i,j)-th subband using the Canny edge detector (Canny, 1986). The algorithm is described as follows in detail.

```
For all the edges of {x(m, n)} do
    For all the edge points (m, n) do
        If (m, n) is an edge point of {x(m, n)} then
            Set the vector containing (m, n) into Λij
            and declare the vector containing
            [<m/ki>,  <n/lj>] to be an element of Λij
        Endif;
        If (m, n) is an element of Λij and
        max{|m′−m|, |n′−n|}≤1 then
            Set the vector containing (m′,n′) into Λij
        Endif;
    Endfor;
Endfor
```

Vector quantization (VQ)

Vector quantization (VQ) is an efficient technique for exploiting the cross-correlation. According to information theory, the bigger the vector dimension, the more correlation removed by the vector quantizer. Traditional vector quantization consists of building a set of signal adapted codevectors called the codebook, with each codevector being assigned a fixed length codeword (signal independent).

A set of well-designed codebooks is crucial for good performance of the reconstruction image. In the frequency domain, the frequency is divided into seven subbands after Walsh transform. Thus, the vectors are classified into seven classes by their subbands in VQ.

A codebook was designed for each sub-region of each subband, with Linde-Buzo-Gray (LBG) algorithm being used to train and obtain the corresponding codebook. For detail of the LBG algorithm, please refer to Linde *et al.*(1980). Each sub-region of each

subband was straightforwardly partitioned into vectors of 4×4 pixels and trained using the LBG algorithm to obtain the corresponding codebook. Full search algorithm was used in the decoder to select the code of the current vector in the corresponding codebook and transmit the address of this code to the channel. The codebook in the decoder was located by the address from the channel and yielded the vector of the code in the corresponding codebook.

In our approach, dynamic rate allocation was not adequate. Our goal was to use vector quantization coding techniques in the subbands, and to rely on the optimal rate allocation to discriminate between all the possible techniques on each sub-region of each sub-band. Different number of bits was assigned to each sub-region of each subband by assuming its distributions. Our bit allocation scheme could allocate more bits for the vectors of the LFS and fewer bits for the vectors of the HFS.

Encoder for LFS and HFS

Different encoders were designed for the LFS and the HFS. It is known that the properties of the LFS are similar to those of the original image and that lower relationship exists in the HFS, so the Vector DPCM (VDPCM) was adopted for encoding the LFS, and Vector Quantization (VQ) coding for encoding the HFS, which will be described in detail.

(1) VDPCM for LFS

Because of the strong relativity of coefficients of the LFS in low and column direction, we selected the VDPCM method to encode the LFS. Not only can their relativity be eliminated, but the efficiency can also be improved because it is the error signal of prediction, which corresponds to the higher frequency of the image.

Two-dimension VDPCM was selected in this algorithm; Fig.4 shows the reconstruction. The predicted value of the current vector can be calculated as the following predictor: $x(i, j)=0.5 \times x(i-1, j)+0.5 \times x(i, j-1)$, where $x(i, j)$ is the current vector of 4×4 pixels.

	$x(i-1,j)$	
$x(i,j-1)$	$x(i,j)$	

Fig.4 Position of reference and current vectors

In the case of the border, one of the coefficients may be absent, then the formula above can be transformed into $x(i, j)=x(i-1, j)$ (if $x(i, j-1)$ is absent) or $x(i, j)=x(i, j-1)$ (if $x(i-1, j)$ is absent). It is noteworthy that, for the first vector of the upper left corner, the coefficients are all absent, so it cannot be predicted but can only be saved in the output code flows of the head information.

In our experiment, the sub-regions of the LFS were directly partitioned into vectors of 4×4 pixels, then the error vectors were trained using the LBG algorithm to obtain the corresponding codebook of the LFS sub-regions.

(2) VQ for HFS

The HFS sub-regions were directly partitioned into vectors of 4×4 pixels, and then directly encoded using VQ. Regarding the details of the VQ algorithm please refer to Sullivan and Baker (1994) and Li *et al.*(1994).

Entropy coding

Conventional Huffman coding was adopted in this completely developed algorithm. Please see Tenenbaum and Augenstein (1981) for the details of the algorithm. Certainly other entropy coding techniques such as run-length coding and arithmetic coding can also be employed here.

EXPERIMENTAL RESULT AND CONCLUSION

In our experiment, we designed a set of local codebooks from a training set of representative images. The training set of images decomposed by two-level filtering was comprised of five 512×512 images (including Lena, Girl, Mand, Aug25r8 and Pentagon shown in Fig.5 and Fig.6) of seven sub-bands after two-level Walsh filtering. Applying the method offered in this paper on Lena and Averser8, we measured by the root mean square error (*RMSE*), peak signal-to-noise ratio (*PSNR*), and compression ratio (*CR*) to determine how good was the quality of the reconstructed image. The experimental result showed that the algorithm reconstructs low bit rate
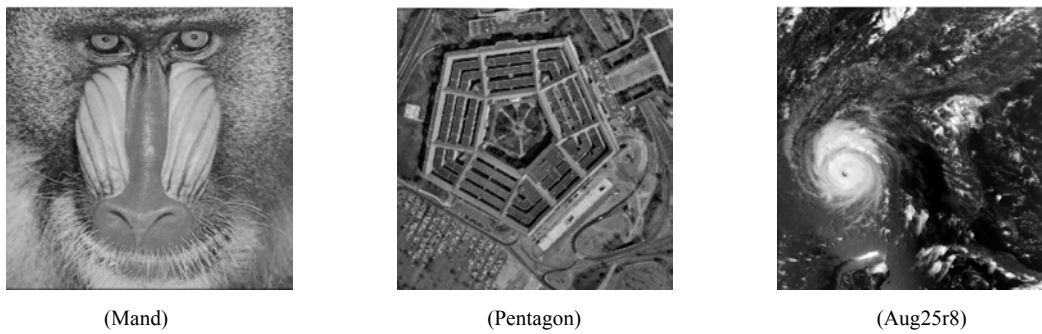

(Girl) (Lena)

Fig.5 Training images in the experiment

(Mand) (Pentagon) (Aug25r8)

Fig.6 Training images in the experiment

Table 1 Results for different images compared with those of JPEG

Algorithm	Image	*RMSE*	*PSNR*	*CR*
Proposed algorithm	Lena	2.493	40.196	9.841
	Averser8	2.716	39.451	8.854
JPEG (Penne, 1988)	Lena	3.3	37.7	8.4
	Averser8	4.5	35.0	5.2
JPEG2000 (ISO/IEC JTC1/SC29/WG1 N1577, 2000)	Lena	2.6	39.7	8.68
	Averser8	3.5	38.4	8.75

coded images with not only high *SNR* but also good perceptual quality, as shown in Table 1. Fig.7 shows the coding result for the Lena and Averser8 image, respectively.

$$MSE = \frac{1}{N} \sum_{i=1}^{N} (x_i - \hat{x}_i)^2$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (x_i - \hat{x}_i)^2}$$

$$PSNR = 10\log_{10} \frac{255^2}{MSE}$$

(Reconstructed Lena)   (Reconstructed Averser8)

Fig.7 Reconstructed images in the experiment

CONCLUSION

In this paper, we present a new algorithm that can yield high fidelity, nearly lossless compression at relatively high compression ratio. Compared with other coding methods, our algorithm has the following several advantages:

1. Our algorithm can yield better reconstruction image quality and higher compression ratio than other traditional algorithms such as JPEG.

2. Compared with typical VQ coding, this algorithm has higher *PSNR* and no block trouble. In particular, there is less mosquito noise in reconstructed images.

3. Compared with traditional subband coding, this algorithm has higher compression ratio and no ringing effect.

The compression performance of our algorithm shows that it can achieve high fidelity image compression with great improvement in compression ratio. So this algorithm has promising prospect in field of high fidelity image compression, especially the compression field of satellite remote sensing images, medical images, and so on.

References

Canny, J., 1986. A Computational Approach to Edge Detection. IEEE Trans. Pattern Anal. Machine Intelligence, PAMI-6, p.679-698.

Ching, P.C., Goodyear, C.C., 1984. Walsh-transform coding of the speech residual in RELP coders. *Proc Inst Elect Eng G*, **131**(1):29-34.

ISO/IEC JTC1/SC29/WG1 N1577, 2000. JPEG2000 Part II Working Draft Version 1.0 Pre-Release A.

Kwon, O.J., Chellappa, R., 1998. Region adaptive subband image coding. *IEEE Trans Image Processing*, **7**:632-648.

Li, W., Wus, J., Zhang, Y.Q., 1994. New Vector Coding Schemes for Image and Video Compression. SPIE Visual Communications and Image Processing, Chicago, IL, p.401-412.

Linde, Y., Buzo, A., Gray, R.M., 1980. An algorithm for vector quantiacer design. *IEEE Trans Communication*, **28**:84-95.

Manz, J.W., 1972. A sequence-ordered fast Walsh transform. *IEEE Trans Audio Elect*, **Au-20**:204-205.

Penne, B.W., 1988. JPEG Technical Specification, Revision 8. Working Document, No. JTC1/SC2/WG10/JPEG-150.

Sullivan, G.J., Baker, R.L., 1994. Efficient quad-tree coding of images and video. *IEEE Trans Image Processing*, **3**:327-331.

Tenenbaum, A.M., Augenstein, M.J., 1981. Data Structures Using Pascal. Prentice-hall, New Jersey, p.275-284.

Tanabe, N., Farvardin, N., 1992. Subband image coding using entropy-coded quantization over noisy channels. *IEEE J Select Areas Commun*, **10**:926-943.