Proceedings IWISP '96; 4-7 November 1996; Manchester, U.K.
B.G. Mertzios and P. Liatsis (Editors)

Visual Vector Quantization For Image Compression Based on Laplacian Pyramid Structure

†Z. He, ‡G. Qiu and †S. Chen
†University of Portsmouth, U.K. ‡University of Derby, U.K.

Abstract

In this paper, we propose a new image coding scheme based on the Laplacian pyramid structure (LPS) and the visual vector quantization (VVQ). In this new scheme, the LPS is used to generate the residual image sequence, and the VVQ is used to code these residual images. Comparing with other block-based coding methods, the new scheme has much less blocking effects on the reconstructed image since coding is performed on the basis of hierarchical multiresolution blocks. The new scheme also has an additional advantage of a much lower computational cost over traditional vector quantization (VQ) techniques since encoding and decoding are based on much smaller dimensional 'visual vectors'. Experimental results show that the new scheme can achieve comparable rate distortion performance to that of traditional VQ techniques, while the computational complexity of the new scheme is only a fraction of that of traditional VQ techniques.

1 Introduction

In recent years, the demand for image transmission and storage has increased dramatically and research into efficient techniques for image compression has attracted extensive interest. Among many coding techniques, the LPS [1] and the VVQ [2] are two efficient coding techniques in terms of compression ratio, fidelity and computational expense.

In this paper, we propose a new image coding scheme by combining the LPS and the VVQ, which inherits the advantages of the both techniques. In this new scheme, the LPS is employed to generate the residual image sequence and the VVQ is used to code these residual images. Experimental results show that the new scheme can achieve comparable rate distortion performance to that of traditional VQ techniques, while the computational cost of the new scheme is much lower since the encoding and deoding are based on much smaller dimensional 'visual vectors'. Because the coding operation is performed on the basis of hierarchical multiresolution blocks, the new scheme has much less blocking effects on the reconstructed image than that of traditional VQ techniques.

The remaining of the paper is organized as follows. Section 2 summarizes the LPS and the VVQ system for coding Laplacian residual images is described in section 3. Section 4 discusses the image reconstruction. Section 5 presents experimental results and section 6 gives some conclusion remarks.

2 The Pyramid Structure

The generation of the pyramid structure includes the generation of the Gaussian pyramid and the generation of the Laplacian pyramid. The process is illustrated in Fig.1.

Gaussian Pyramid Generation

The original image G_0 of size $M \times N$ pixels becomes the level 0 of the Gaussian Pyramid. Upper level images are generated by using the reduction function $R(\cdot)$ [1] defined in (1), iteratively.

$$G_l(i,j) = \sum_{m=-2}^{2} \sum_{n=-2}^{2} w(m,n) \cdot G_{l-1}(2i+m, 2j+n) \qquad 0 < l \leq L, \ \ 0 \leq i < M_l, \ \ 0 \leq j < N_l. \quad (1)$$

L is the number of levels in the pyramid, M_l and N_l are the dimensions of the lth level, and $w(m,n)$ are weighting kernels. Fig.2 shows a 5-level Gaussian pyramid of "Lena".

Laplacian Pyramid Generation

The reverse of the reduction function $R(\cdot)$ is the expansion function $E(\cdot)$ [1] defined in (2). Let the $G_{l,n}$ be the result of expanding G_l n times. Then

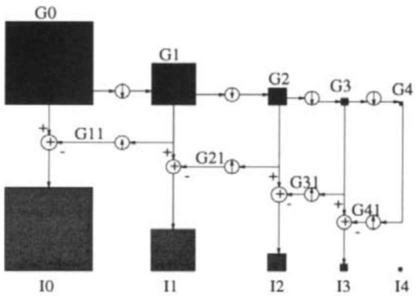

Figure 1: Pyramid Structure Generation

Figure 2: 5-Level Gaussian Pyramid of "Lena"

$$G_{l,n}(i,j) = 4\sum_{m=-2}^{2}\sum_{n=-2}^{2} w(m,n)\cdot G_{l,n-1}\left(\frac{i-m}{2},\frac{j-n}{2}\right) \qquad 0 < l \leq L,\ \ 0 \leq n \leq l, \qquad (2)$$
$$0 \leq i < M_{l-n},\ \ 0 \leq j < N_{l-n}$$

The Laplacian pyramid is a sequence of residual images $I_0, I_1, \cdots, I_L$, each being the difference of two adjacent levels of the Gaussian pyramid. Thus

$$\begin{aligned} I_l &= G_l - G_{l+1,1} \qquad for\ \ 0 \leq l \leq L-1 \\ I_L &= G_L \end{aligned} \qquad (3)$$

Fig.3 shows a 5-level Laplacian pyramid of "Lena" image generated by Eqn.(3).

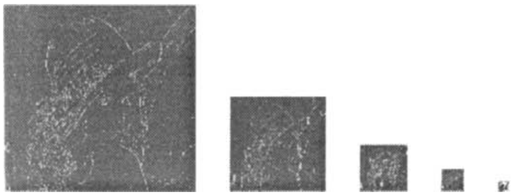

Figure 3: 5-Level Laplacian Pyramid of "Lena"

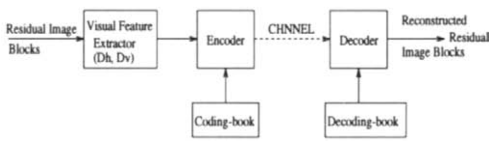

Figure 4: VVQ Image Coding System

3 Visual Vector Quantization

The design of the VVQ coding system consists of the design of the coding-book used in coding phase and the design of the decoding-book used in decoding phase.

Design of Coding-book

The residual image I_l of size $M_l \times N_l$ is divided into $P_l \times Q_l$ blocks of size $m_l \times n_l$, where $P_l = M_l/m_l$, $Q_l = N_l/n_l$. The horizontal and vertical derivatives [2] of each block are calculated as

$$D_h(p,q) = \frac{1}{8}\sum_{i=0}^{3}\sum_{j=0}^{3} I_l(4p+i, 4q+j)\cdot g_h(i,j) \qquad 0 \leq p < P_l,\quad 0 \leq q < Q_l \qquad (4)$$

$$D_v(p,q) = \frac{1}{8}\sum_{i=0}^{3}\sum_{j=0}^{3} I_l(4p+i, 4q+j)\cdot g_v(i,j) \qquad 0 \leq p < P_l,\quad 0 \leq q < Q_l \qquad (5)$$

where the values of the kernels g_h and g_v can be written collectively in matrix form as

$$G_h = \begin{bmatrix} 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 \end{bmatrix} \qquad G_v = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ -1 & -1 & -1 & -1 \\ -1 & -1 & -1 & -1 \end{bmatrix}$$

The horizontal and vertical derivatives of a block are used to form a "visual vector", (D_h, D_v), to represent the block. The visual vectors of all blocks of residual image I_l are partitioned into N_c clusters using the competitive learning [3], and these cluster centers are used to comprise the coding-book for residual image I_l.

Design of Decoding-book
A multilayer neural network is trained using backpropagation [4] to reproduce the residual blocks of I_l when the corresponding visual vectors are presented at the input. The decoding-book is obtained from the output of the trained network by feeding the cluster centers to the input. The network has 2 input neurons and $m_l \times n_l$ output neurons. The number of the hidden-layer neurons is decided by experiment. The whole VVQ image coding system is illustrated in Fig.4.

4 Image Reconstruction

The reconstruction of the original image $\hat{G}_0$ from the decoded residual image sequence $\hat{I}_0, \hat{I}_1, \cdots, \hat{I}_L$ is achieved by reversing the operations of the Laplacian pyramid generation as follows:

$$\begin{array}{ll} \hat{G}_L = \hat{I}_L & \hat{G}_{L,1} = E(\hat{G}_L) \\ \hat{G}_{L-1} = \hat{G}_{L,1} + \hat{I}_{L-1} & \hat{G}_{L-1,1} = E(\hat{G}_{L-1}) \\ \cdots\cdots & \cdots\cdots \\ \hat{G}_1 = \hat{G}_{2,1} + \hat{I}_1 & \hat{G}_{1,1} = E(\hat{G}_1) \\ \hat{G}_0 = \hat{G}_{1,1} + \hat{I}_0 & \end{array} \tag{6}$$

where $E(\cdot)$ is defined as in (2).

5 Experimental Results

Two 512×512 monochrome images, "Lena" and "Peppers", with 8 bits per pixel, were used to evaluate the proposed new scheme. The "Lena" image was used to train the system and the "Peppers" image was used to test the system.

A 5-level pyramid structure and a 4-level pyramid structure were investigated separately. The highest level residual image of the pyramid structure was coded directly using 8 bits per pixel. All other lower level residual images were coded using VVQ with the 4×4 block size for levels 0 and 1 and the 2×2 block size for levels 2 and 3, respectively. The coding-book size was chosen as $N_c = 9$. A 3-layer neural network with 2 input neurons, 10 hidden neurons and 16 or 4 output neurons was used to generate the decoding-book, depending on whether the 4×4 or 2×2 block size was used.

In the test, it was found that a large amount of blocks fell into the class which had least significant edge contents. Hence a variable bit rate coding strategy was used. One bit was used to code the blocks which belong to this class, and 4 bits for each of the other 8 classes. Fig.5 and Fig.7 are the images reconstructed by the proposed coding scheme using 5-level and 4-level pyramids, respectively. As a comparison, Fig.6 and Fig.8 show the images reconstructed by the traditional VQ technique LBG [5] using the codebooks of size 8 and 16, respectively, with a block size of 4×4. The performance and the computational cost of the proposed new scheme and the LBG are summarized in Table 1.

Experimental results show that the proposed new scheme has much less blocking effects on the reconstructed image compared with the conventional VQ technique. This results in a smother reconstructed image as is evident in Fig.5-8. From Table 1, it can be seen that, at similar bit rate and peak signal to noise ratio (PSNR), the computational requirements of the new scheme is only a small fraction of those of the LBG scheme.

6 Conclusions

A new image coding scheme has been proposed based on a combined LPS and VVQ approach. Since the coding is performed on the basis of hierarchical multiresolution blocks, the proposed new scheme has much less blocking effects on the reconstructed image compared with other block-based techniques. The computational cost of the proposed scheme is also much lower than traditional VQ techniques because the new scheme uses much smaller dimensional visual vectors to represent image blocks.

References

[1] P. J. Burt and E. H. Adelson, " The Laplacian pyramid as a compact image code," *IEEE Trans. Comm.*, Vol. COM-31, No. 4, pp532-540, April 1983.

[2] G. Qiu, M. R. Varley and T. J. Terrell, "Image coding based on visual vector quantization," *Image Processing and Its Applications, IEE Conference Publication* No.410, pp301-305, July 1995

[3] T. Kohonen, *Self-Organization and Associative Memory.* Second Edition, Berlin, Springer-Verlag, 1988.

[4] D. E. Rumelhart, G. E. Hinton and R. J. Williams, "Learning internal representations by error propagation," Chapter 8 in *Parallel Distributed Processing,* Vol. 1, Cambridge, MA, MIT Press, 1986.

[5] Y. Linde, A. Buzo and R. Gray, "An algorithm for vector quantizer design," *IEEE Trans. Comm.,* Vol. COM-28, No. 1, pp84-95, January 1980.

Table 1: Comparison of Proposed Scheme and LBG

Coding Scheme		Performance		Computational Cost	
Scheme	Parameters	bitrate(bpp)	PSNR(dB)	Addition No.	Multiplication No.
Proposed	5-L Pyramid	0.18	24.71	153,600	25,600
LBG	codebook size 8	0.19	25.60	8,388,608	8,388,608
Proposed	4-L Pyramid	0.27	27.49	137,216	24,576
LBG	codebook size 16	0.25	26.63	16,777,216	16,777,216

Figure 5: New scheme reconstruction (5-L pyramid)

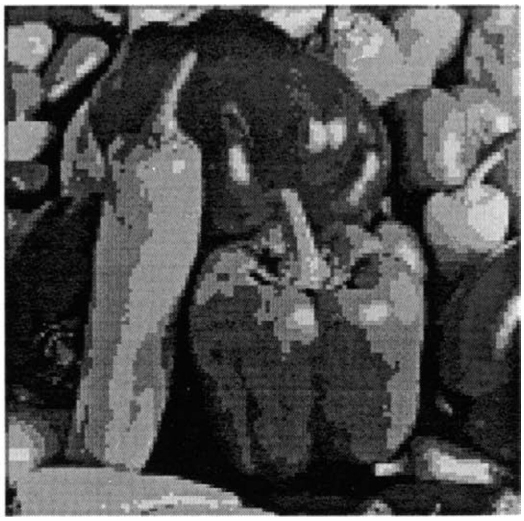

Figure 6: LBG reconstruction (codebooksize 8)

Figure 7: New scheme reconstruction (4-L pyramid)

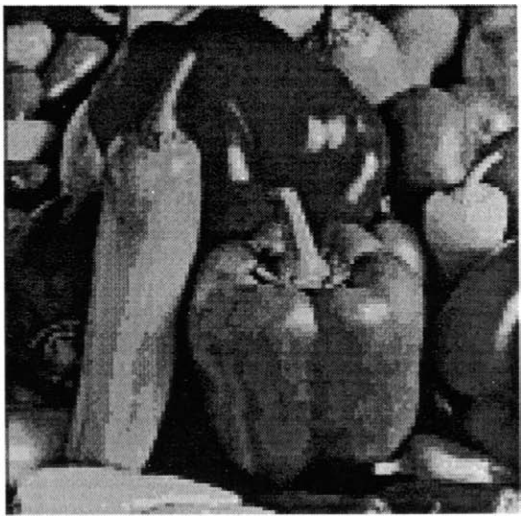

Figure 8: LBG reconstruction(codebooksize 16)