Classified Wavelet Transform Coding of Images Using Vector Quantization

Y. Huh, J. J. Hwang, C. K. Choi, R. de Queiroz and K. R. Rao
Electrical Engineering Department
University of Texas at Arlington
Box 19016, Arlington, TX, 76019

Abstract

The discret wavelet transform (DWT) has recently emerged as a powerful technique for image compression in conjunction with a variety of quantization schemes. In this paper, a new image coding scheme - classified wavelet transform/ vector quantization (DWT/CVQ) - is proposed to efficiently exploit correlation among different DWT layers aiming to improve its performance. In this scheme, DWT coefficients are rearranged to form the small blocks, which are composed of the corresponding coefficients from all the subbands. The block matrices are classified into four classes depending on the directional activities, i.e., energy distribution along each direction. These are further divided adaptively into subvectors depending on the DWT coefficient statistics as this allows efficient distribution of bits. The subvectors are then vector quantized. Simulation results show that under this technique the reconstruction images preserve the detail and structure in a subjective sense compared to other approaches at a bit rate of 0.3 bit/pel.

Keywords: discret wavelet transform, transform coding, classified vector quantization

1 Introduction

Transform coding(TC) of images has proved to be an efficient compression technique and has been applied in various forms [1]. According to Shannon's rate distortion theory, vector quantization(VQ) of signals reduces the coding bit rate significantly compared to scalar quantization [2, 3]. Especially, classified transform VQ (CVQ) reduces edge degradation and improves coding efficiency [4]. This technique also takes advantage of the decorrelation and energy compaction properties of transform coding and the superior rate distortion performance of VQ [5]. Classification in transform domain results in efficient bit assignment with only a slight increase in bit rate due to overhead information [6].

The DWT has recently emerged as a powerful technique for image compression because of its flexibility in representing images and its ability in adapting to the human visual system characteristics. Several compression techniques using the wavelet transform and different types of quantization methods have been proposed in the literature [7, 8, 9, 10, 11]. It decomposes an image into various multi-resolution approximations, which are accomplished by iteratively applying high- and low-pass filters to the image. The advantage of using the DWT over the discrete cosine transform (DCT) lies on the fact that the DWT projects high-detail image components onto shorter basis functions, with higher resolution, while lower detail components are projected onto larger basis functions, which corresponds to narrower subbands, establishing a trade between time and frequency resolutions [12]. In addition, the wavelet transform coding provides a superior image quality at low bit rates, since it is free from both blocking effects and mosquito noise [13]. However, there are still significant amounts of redundancies among the subbands. That is, most coding techniques using the DWT disregard cross-correlation among the various scale and orientation bands. Hence, the performance of the wavelet transform coding can be improved if the cross- correlations among these subbands are exploited in the encoding process.

In this paper, we propose a new coding scheme for image compression using classified VQ of the wavelet coefficients, which exploits the residual correlation between different layers and improves the encoding efficiency by taking advantage of DWT and CVQ. In this scheme, DWT coefficients are rearranged to form the small blocks, which are composed of the corresponding coefficients from all the subbands. The block matrices are classified into four classes depending on the directional activities and then each subvector of each class is quantized and coded using its own codebook, which is generated for each subvector from a training sequence. In section 2, a brief review of

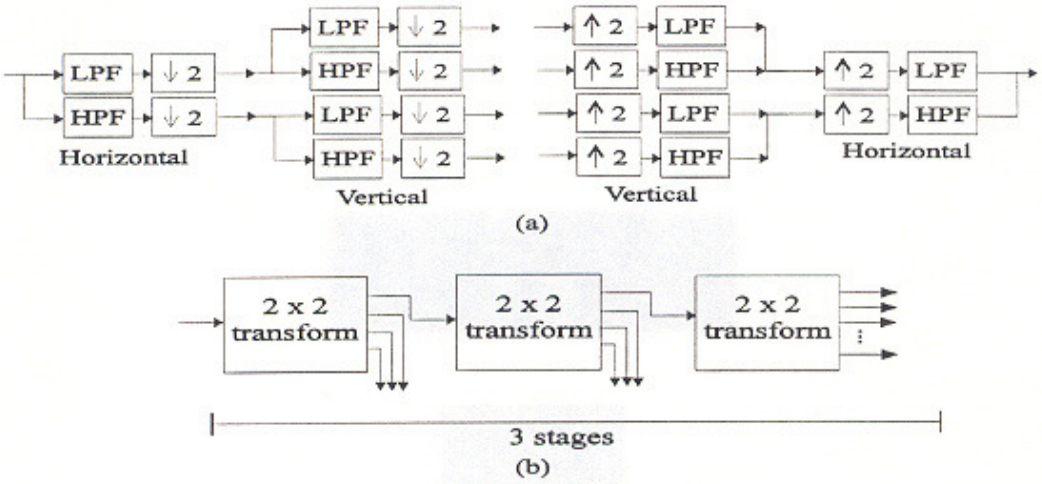


Figure 1: Decomposition and reconstruction of the DWT.

DWT decomposition is described. Section 3 discusses classified VQ in the DWT domain for exploiting the residual correlations. The improved performance compared to other coders is presented in section 4.

2 Discrete wavelet transform

The DWT is known to be generated by a cascade of filter banks and it is essentially based on the well-known subband decomposition. However, in its most popular form, the dyadic DWT, the input spectrum is partitioned into octave-width subbands. The advantage comes from the trade-off between spatial and frequency resolutions, as the DWT has shorter basis functions (filters) for higher frequencies, and longer basis functions (filters) for lower frequencies. Also, there are more samples to represent the higher frequency subbands, than the lower frequency ones. Therefore, more samples and shorter basis functions will attain a better subband selectivity, for lower frequencies, and as low frequency components lack details, spatial resolution is less important,in this case.

Fig. 1(a) shows the analysis section of a two-dimemsional(2D) seperable filter bank, where first the image rows are passed through the 2-channel filter bank, and then the columms are processed. The right side of Fig. 1(a) shows the synthesis section to reconstruct the signal from the subband signals. The analysis section can be viewed as a 2×2 transform applied to the image (note that each subband has one fourth of the samples in the original signal). Also, the synthesis section can be viewed as a 2×2 inverse transform. Fig. 1(b) shows the scheme we intend to apply, which is composed of a succession of 3 stages of 2×2 transforms. The inverse transform is, of course, accomplished by reversing the paths and the transforms. The decomposition of Lena image is shown Fig. 2 using these schemes. To reconstruct the image from the subbands, one may choose the set of low-pass and high-pass filters in Fig. 1(a), as to provide perfect reconstruction(PR). However, PR 2-channel filter banks, either have non-linear-phase or are not orthogonal. As we intend to apply the DWT to low-bit-rate compression we decided to use Johnston's near-PR filters [14], which have linear-phase and very small reconstruction error. Also they can be easily implemented using symmetric extensions [15], being free of any border distortions.

3 Classified DWT/VQ

The CVQ is well understood in terms of the composite source model for images, where the image is viewed as a bank of subsources [4]. To apply CVQ in the DWT domain, DWT coefficients are rearranged to form the small blocks, which are composed of the corresponding coefficients from all the subbands. The design of a block classifier that effectively classifies the input block (rearranged small block) according to its activity is important. In order to get a good quality of the reconstructed image at relatively low bit rates, the quantization has to adapt to the structure in the DWT domain. Thus, the product code VQ scheme that partitions the DWT block into several smaller subvectors

Figure 2: Decomposition of Lena image.

and performs VQ on these subvectors is employed. In this case the selection of an efficient partitioning scheme is also important. The entire coding scheme is shown in Fig. 3.

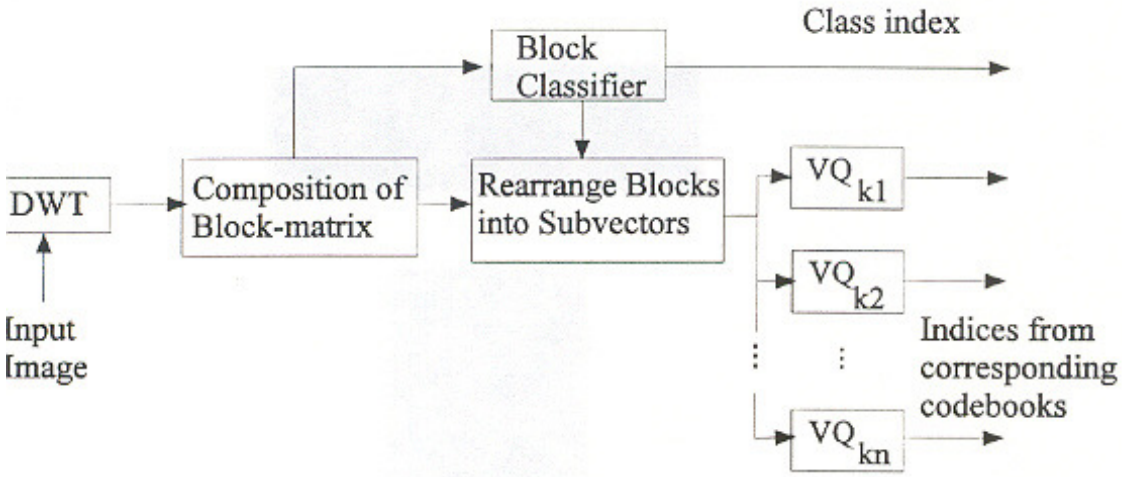
3.1 Block rearrangement

The input image is first decomposed into a pyramid structure with 3 levels by using the DWT. In order to reduce the correlation among the different layers, transform coefficients are rearranged to form the small blocks, which are composed of the corresponding coefficients from all the subbands as shown Fig. 4. The size of small block is $N \times N$, where $N = 2^{(\#\text{ of levels})}$. In our scheme, therefore, the 8×8 matrix for a small block was constructed as the (8×8) DCT has been most universally applied in image coding. Here, the size of subblock is $2^{(3-l)}$ (l = 1,2,3) for each horizontal, vertical, and diagonal orientation band at the l-th level. Thus each 8×8 matrix has three subblocks of size 4×4 in level I, three subblocks of size 2×2 in level II, and four elements in level III.

3.2 Block classification

The main purpose of block classification is to separate the blocks into perceptually distinct categories (overhead bits are needed to indicate the categories of the blocks) and to use a different quantization scheme for each category. Blocks from different activity classes are then independently quantized with a bit allocation appropriate for each class. The block matrices are classified into four classes depending on the directionl activities, i.e., ac energy distribution along each direction. Though, the directional activities are essentially based on the subband distribution in wavelet transform domain, final decisions are made by combining all subbands in a block. A similar scheme [5, 6] using the activity measure defined in DCT domain is employed in this DWT/CVQ.

But, we propose an adaptive threshold decision technique to classify the four classes of the image. A block is, first, classified into two classes: low or high activity one. The decision criterion is to set an average directional energy of all training images. No special values are applied for the test images. We assume that the directional energy distributions are Gaussian with non-zero mean which we use for the first classification. The second threshold is set as half of mean energy and the third one is symmetrically defined. The lowest energy region is, of course, classified as non-active (uniform). The second region, however, must be classified as directional, if any activity is relatively



igure 3: Block diagram for classified wavelet transform/VQ coding of images.

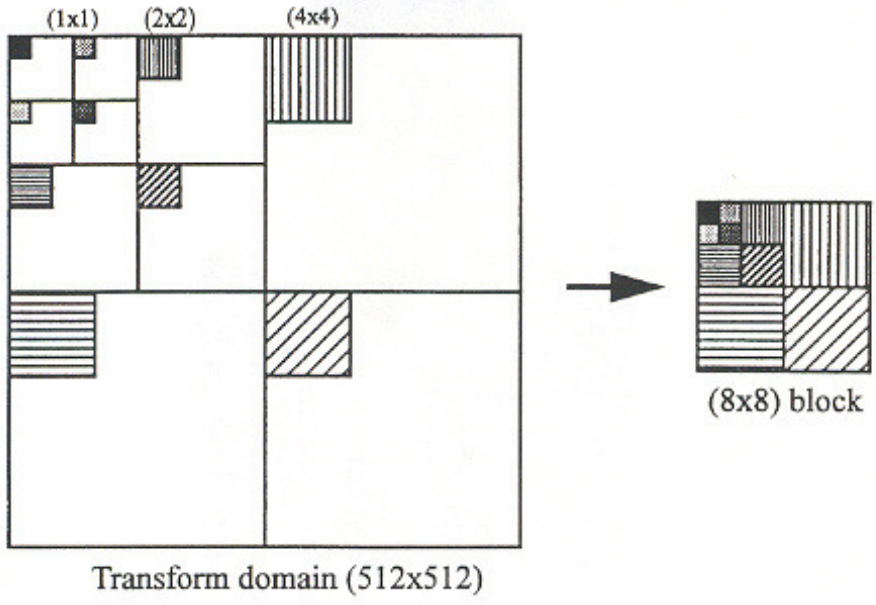


Figure 4: Composition of the block matrix (3 levels).

Table 1: Typical variance distribution for horizontal class in the (8×8) DWT domain. Subvector configurations corresponding to this variance distribution are shown in Fig. 5.

2267.939	243.460	111.941	113.038	27.306	30.937	30.476	28.838
9.974	26.466	117.630	113.583	31.468	31.084	33.270	32.874
7.300	7.064	11.083	10.084	29.198	34.216	33.212	30.669
8.489	7.262	10.454	11.180	27.347	28.457	29.164	28.625
5.032	4.241	4.900	4.296	3.112	3.108	3.487	3.332
4.022	4.422	4.633	4.370	2.869	3.491	3.360	3.123
4.373	4.272	4.513	4.240	2.831	3.433	3.886	3.340
4.718	4.628	4.360	4.325	3.094	2.892	3.368	3.258

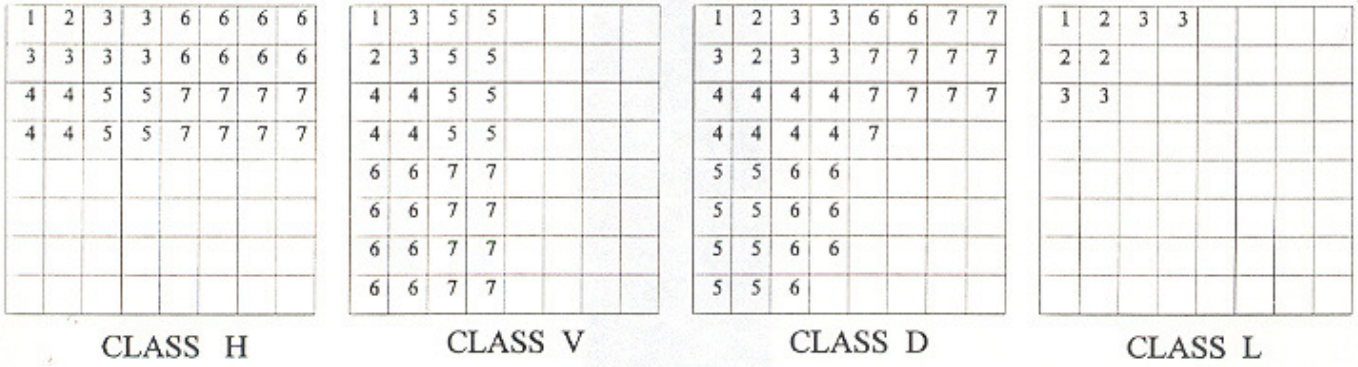


Figure 5: Subvector configurations. Coefficients with the same digit belong to a subvector

large (The human visual system is quite sensitive to edges i.e., horizantal, vertical, or diagonal). Otherwise, it is classified as non-active (or low active).

According to the contrast sensitivity of the HVS, the mininum luminance difference is proportional to the brightness of the background. In our coder, the background levels can be block averages (DC coefficients), which are easily obtained for the smaller blocks. In the third class, the directional energies are compared and the largest one can be a directional moment in the block. If three directional energies are greater than the third threshold, then the block is classified as non-directional (diagonal). This adaptive technique can classify the input image efficiently because no predefined thereshold value for each image is needed.

3.3 Subvector construction

The blocks are partitioned into a number of subvectors(1-7) based on the variance distribution of the coefficients in each class. The results of such partitioning is a product code. The partition chosen for each of the four classes are shown in Fig. 5. These partitions are determined in such a manner that coefficients that have similar variances are grouped into a single subvector. For example, the variance map for the horizontal class is shown in Table 1. In this case we can easily construct subvectors due to the nature of the DWT as shown in Fig. 5. Some orientation bands in this block were grouped together so that the interband correlation is efficiently utilized.

Large number of bits are assigned to subvectors which have coefficients with large variances and proportionally fewer bits to those with small variances. Hence, bits are allocated according to the total variance of each subvector. Since most of the energy is concentrated in a few coefficients, some of the coefficients with small energy levels can be excluded from the coding process. To achieve the high compression ratio, most high frequency coefficients are discarded as well. Each subvector of a class has a unique codebook since the statistics and dimensions of subvectors are different. Hence, the effective bit rate reduction is achieved and the problem of high computational complexity of conventional VQ is alleviated. An example for the bit allocation and corresponding distortion (mse) for each class for overall average bit rate at 0.5 bit/pel (bpp) is shown Table 2.

Table 2: Bit allocation of DWT/CVQ for 0.5 bpp rate scheme

Subvector	Class H			Class V			Class D			Class L		
	dim.	bit	dis.	dim.	bit	dis.	dim.	bit	dis.	dim.	bit	dis.
1	1	8	-	1	8	-	1	8	-	1	8	-
2	1	8	0.018	1	7	0.050	2	8	0.700	3	6	0.512
3	6	10	4.502	2	7	0.727	5	9	6.918	4	5	0.629
4	4	7	1.179	4	8	6.068	8	9	6.652			
5	4	5	2.915	8	7	4.426	8	4	8.538			
6	8	8	6.377	8	7	3.962	9	7	5.490			
7	8	8	6.422	8	6	4.723	11	8	7.091			

* dim: subvector dimension, bit: codebook size = 2^{bit}

Table 3: DWT/CVQ simulation results for several bit rates

Encoded image	Bit rate (peak signal to noise ratio)				
Lena (512×512)	0.20 bpp (29.44 dB)	0.21 bpp (30.21 dB)	0.33 bpp (32.27 dB)	0.47 bpp (33.02 dB)	0.50 bpp (34.11 dB)
Boat (512×512)	0.21 bpp (30.39 dB)	0.31 bpp (31.48 dB)	0.33 bpp (32.73 dB)	0.47 bpp (34.27 dB)	0.50 bpp (34.42 dB)

3.4 Overhead information

The class information for each block is included in the coded image so that reconstruction is possible. This overhead information indicates to the decoder which codebook set has to be used and also the way in which the subvectors have to be recombined to form the entire 8×8 block. Let there be B subblocks in an image. The number of bits to represent the K classes is $\log_2 K$. Thus the overhead bits for classification are $b_0 = B \log_2 K$. For 512×512 - pels images, and (8×8) block size, the overhead information for four classes is $b_0 = 4096 \times 2 = 8192$ bits. The total number of overhead bits is 8192 bits or, equivalently, 0.031 bpp for four classes.

4 Simulation results

Coding simulations were performed for monochrome images of 521×512 - pels, originally with 8 bpp. Two test images, 'Lena'and 'Boat', are shown in Fig. 6. Test images were coded using the codebook generated from a number of training sets. A total of four images were used in the training sequence to generate the VQ codebooks.

After classifying the block into four classes and partitioning into subvectors, the dc coefficients are quantized using the scalar quantization. All the ac coefficients are quantized using the codebooks generated by the LBG algorithm [17]. The bit allocation for each of subvectors and total number of bits depend on the target output bit rates. For a given quota of bits, therefore, we control the number of bits for each subvector of each class based on the distortion error rate. As an objective measure of reconstructed image quality, we used the peak signal-to-noise ratio (PSNR) as

$$PSNR \equiv 10 \log_{10} \frac{255^2}{MSE} \tag{1}$$

where

$$MSE = \frac{1}{M^2} \sum_{i=1}^{M} \sum_{j=1}^{M} [x_{ij} - y_{ij}]^2 \tag{2}$$

Table 4: Comparison of PSNR values among different coding schemes

Image	Bit rate (bpp)	PSNR(dB)		
		DWT/CVQ	LOT/CVQ	DCT/CVQ
Lena (512×512)	0.33	32.3	30.6	29.6
	0.5	34.1	33.7	32.6

where x_{ij} and y_{ij} are the (i,j) th pixels in the original and reconstructed image, respectively, and the images are of size $M \times M$. The result of the several bit rates for two test images are shown in Table 3. Even at 0.21 bpp, our coder yields 30.21 dB for image Lena. Three reconstructed Lena and Boat images at 0.33, and 0.5 bpp (including overhead bits) are also shown in Figs. 7 and 8, respectively. No blocking effects as in DCT based coder are perceptible in our proposed scheme at a bit rate of 0.33 bpp. As can be seen, the proposed scheme can yield higher compression ratio, while maintaining a good reconstruction quality (both objective and subjective).

A similar coding scheme using DCT and LOT was introduced in [6, 16]. For comparision, the performance of LOT/CVQ based coder proposed in [16] is also shown in Table 4. It is evident from this table that the PSNR performance of our wavelet coder is better than the coder in [16] at all encoding rate.

5 Conclusions

In this paper, we have presented a new image coding scheme using CVQ in the DWT domain. This scheme exploits the residual correlation among different layers of the DWT domain using block rearrangement to improve the coding efficiency. Further improvement can also be made by developing the adaptive threshold techniques for classification based on the contrast sensitivity characteristics of the human visual system. The coding complexity of DWT/CVQ is less than that of a ordinary VQ because of the decomposition of the transform block into subvectors. Simulation results have shown that the coder employing DWT/CVQ outperforms the conventional transform coding techniques such as DCT/CVQ and LOT/CVQ. One deficiency of the work presented here is the lack of adaptive bit allocation procedure based on the rate distortion bound for various wavelet transform coefficients of an image. We plan to study the bit allocation problem encountered here, devise an appropriate bit allocation strategy for this coder, and report the results in the future.

References

[1] R. J. Clarke, "*Transform coding of images*," Academic Press, New York, 1985.

[2] R. M. Gray, "Vector quantization," *IEEE ASSP Magazine*, vol.1, pp. 4–29, Apr. 1984.

[3] N. Nasrabadi and R. King, "Image coding using vector quantization: A review," *IEEE Trans. on Communications*, vol. COM-36, pp. 957–971, Aug. 1988.

[4] B. Ramamurthi and A. Gersho, "Classified Vector Quantization of images ," *IEEE Trans. on Communications*, vol. COM-34, pp. 1105–1115, Nov. 1986.

[5] Y. S. Ho and A. Gersho, "Classified transform coding of images using Vector Quantization ," *Proc. ICASSP-89*, pp. 1890–1893, Glasgow, Scottland, May 1989.

[6] J. Y. Nam and K. R. Rao, "Image coding using a classified DCT/VQ based on two channel conjugate Vector Quantization," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 1, pp. 325-336, Dec. 1991.

[7] M. Antonini, M. Barlaud, P. Mathieu and I. Daubechies, "Image coding using Wavelet Transform," *IEEE Trans. on Image Processing*, vol. 1, pp. 205–220, Apr. 1992.

[8] P. Sriram and M. W. Marcellin, " Image coding using wavelet transform and entropy-constrained trellis coded quantization," *Proc. ICASSP-93*, pp. V-554–V-557, Minneapolis, MN, Apr. 1993.

[9] Y. Kato, Y. Yamada, H. Ohira and T. Murakami " A vector quantization of wavelet coefficients for super high definition image," *Proc. SPIE/VCIP'93*, vol.2094, pp. 1021–1028, Nov. 1993.

10] N. Moayeri, I. Daubechies, Q. Song and H. S. Wang, "Wevelet transform image coding using trellis coded vector quantization," *Proc. ICASSP-92*, pp. IV-405–IV-408, San Francisco, CA, Mar. 1992.

11] X. Wang and S. Panchanathan, "Wavelet transform coding using NIVQ," *Proc. SPIE/VCIP'93*, vol. 2094, pp. 999-1009, Nov. 1993.

12] M. Vetterli and C. Herley, "Wavelets and filter banks," *IEEE Trans. on Signal Processing*, vol. 40, pp. 2207-2232, Sep. 1992.

13] M. Ohta, M. Yano and T. Nishitani, "Wavelet picture coding with transform coding approach," *IEICE Trans. Fundamentals*, vol. E75-A, pp. 776–785, July 1992.

14] J. D. Johnston, "A filter family designed for use in quadrature mirror filter banks," *Proc. of Intl. Conf. on Acoust., Speech, Signal Processing*, Denver, CO, pp. 291–294, 1980.

15] M. J. Smith and S. L. Eddins, "Analysis/synthesis techniques for subband image coding," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, pp. 1446–1456, Aug. 1990.

16] S. Venkatraman, J. Y. Nam and K. R. Rao, "Image coding based on classified lapped orthogonal transform - vector quantization," *Proc. SPIE/VCIP'92*, vol.1818, pp. 500–511, Nov. 1992.

17] Y. Linde, A. Buzo and R. Gray, "An algorithm for vector quantizer design," *IEEE Trans. on Communications* , vol. COM-28, pp. 84–95, Jan. 1980.

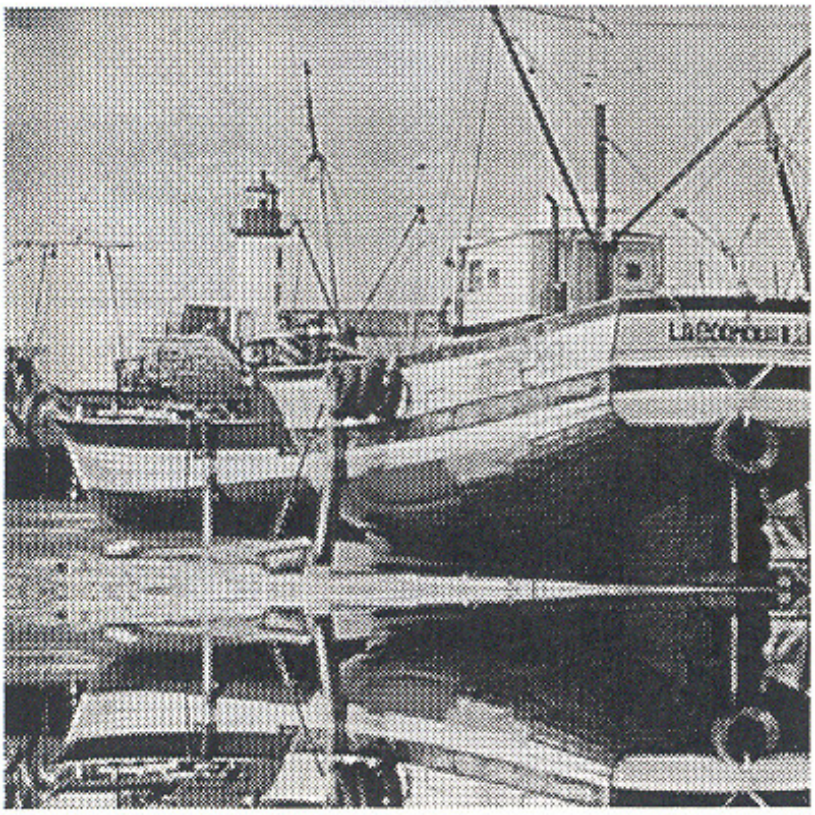
18] T. Omachi, Y. Takashima and H. Okada, "DCT-VQ coding scheme using categorization with adaptive band partition," in *PCS-87 Program & Abstract*, pp.161–162, June 1987.

19] A. Gersho, "Asymptotically optimal block quantization," *IEEE Trans. on Information Theory* , vol. IT-25, pp. 373–380, July 1979.

20] J. W. Kim and S. U. Lee, "A transform domain classified vector quantization for image coding," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 2, pp. 3-14, Mar. 1992.

21] S. G. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Trans. on Pattern Anal. and Mach. Intel.*, vol. 11, pp. 674–693, July 1989.

(a)



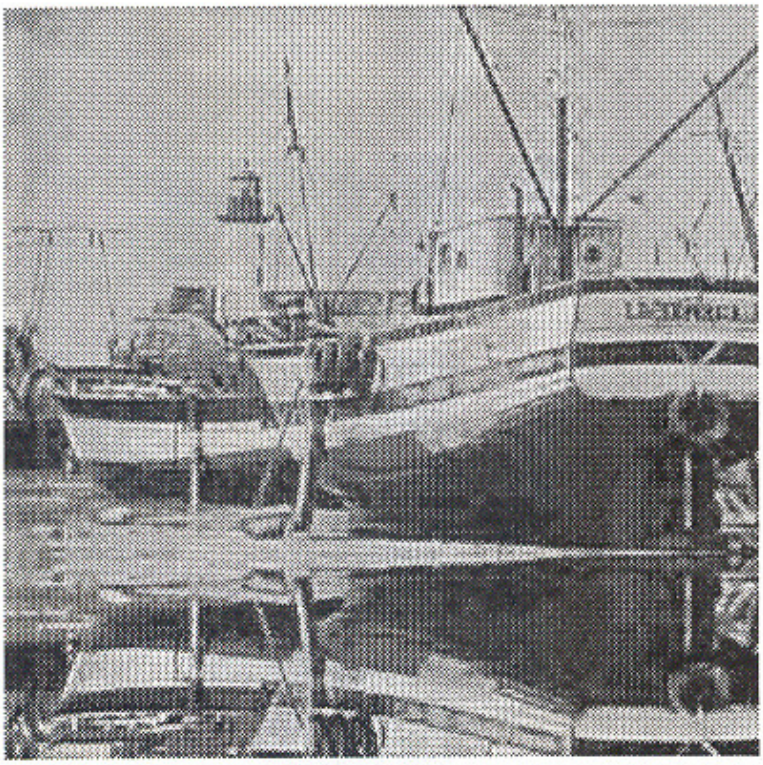
(b)

Figure 6: Original images (a) Lena and (b) Boat

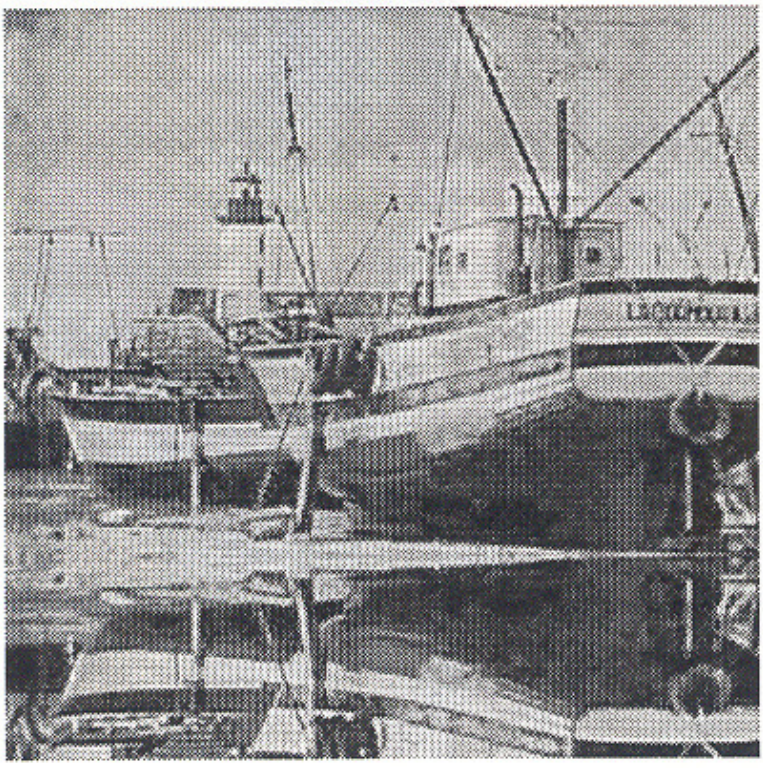
(a)

(b)

Figure 7: Reconstructed Lena image (a) at 0.33 bpp and (b) at 0.5 bpp



(a)



(b)

Figure 8: Reconstructed Boat image (a) at 0.33 bpp and (b) at 0.5 bpp