# A New Hybrid Coder for High Quality Image Compression

**Hang-Chan Lee**

## Abstract

This paper presents a new design technique for performing high quality low bit rate image compression. A *hybrid coder(HC)* which combines *Mean Removed Important Coefficient Selection based JPEG(MR-ICS-JPEG)* and *Adaptive Vector Quantization(AVQ)* is proposed. A new quantization table is developed using the *Important Coefficient Selection(ICS)* method; the importance of each coefficient is determined using the orthonormal property of the DCT. This quantization table is applied to standard *JPEG* with mean removal*(MR)* strategy before processing. This scheme, called *MR-ICS-JPEG*, produces more than 2 dB enhanced performance in terms of PSNR over standard *JPEG*. A set of homogeneous codebooks is generated by homogeneous training vectors. Before compression, an image is uniformly divided into 8×8 blocks. Low detail regions such as backgrounds are roughly coded by *AVQ* while high detail regions such as edges or curves are finely coded by the proposed *MR-ICS-JPEG*. This *hybrid coder* produces consistently about 3 dB improved performance in terms of PSNR over standard *JPEG*.

## I. Introduction

*Vector quantization(VQ)* attempts to code groups of parameters together[1]. As a result, a large dimensionality produces a high compression ratio while resulting in an exponentially increasing number of calculations. Because of this encoding complexity, there are some limitations for employing high dimensional codebooks which have a large number of codewords. If a codebook does not include enough codewords to adequately represent edges or curves, a VQ system cannot avoid producing edge degradation. This limitation restricts VQ for widespread use when it is compared to a DCT based transform coder such as *JPEG*.

A DCT based transform coder uses a fundamentally different coding scheme than VQ in the sense that it is based not on a vector space but a scalar space. Because each DCT coefficient is coded and transmitted separately, a high quality compressed image can be produced while being somewhat limited in achieving high compression. Moreover, attempting to achieve a high compression ratio results in edge blurring because most of the high frequency components are removed by the quantization procedure.

In order to alleviate the inherent problems of VQ and DCT based transform coders listed above, a hybrid coder(HC) which combines the advantage of vector quantization(VQ), high compression, and the advantage of DCT based transform coder (TC), high quality, is presented in this paper.

The hybrid coder can be considered as a synthetic compression technique which combines somewhat complementary compression schemes, VQ and DCT based TC. Usually, the human eye is less sensitive to the low detail regions such as backgrounds while more sensitive to the high detail region such as edges or curves. Therefore, each image block needs to be treated separately according to its activity. In other words, low detail regions are coded roughly while high detail regions are coded finely to represent edges or curves more accurately. This is the main strategy of the hybrid coder for improving image quality.

Usual hybrid schemes employ classified VQ based on edge classifications[3, 4, 5, 6]. However, these schemes produce many overhead bits, often more than 5 bits per block, because many edge patterns exist. Moreover, it is very hard to design an efficient classifier. Because of the reasons listed above, the performance of a classified vector quantizer is usually not outstanding. In order to design a hybrid coder which provides low overhead and a simple classifier, a homogeneity test can be employed.

Before compression, an input image is equally divided into 8x8 blocks. Instead of classifying image blocks according to their edge patterns, we classify image blocks into 8 classes according to their homogeneity. Next, seven classes of image blocks are coded by corresponding homogeneous codebooks where each codebook is generated from different homogeneous training vectors. The remaining class of image blocks, which consists of blocks of high activity, is coded by the TC rather than the VQ. Some addi-

---

Manuscript received August 29, 1997; accepted November 14, 1997.

The Author is with Information Processing Sector at SAIT in Korea.

tional bits which are made available by the codebook adaptive VQ(AVQ) are allocated to the DCT-based TC to represent the high detail regions more accurately. The TC is implemented by a modified *JPEG* scheme, *Mean Removed-Important Coefficient Selection- JPEG (MR-ICS-JPEG).*

Since eight classes are used, resulting overhead bits are restricted to no more than 3 bits for each 8x8 image block in the HC. The adaptivity which allows higher bit allocation to high detail regions, while requiring low overhead, is the main reasons the proposed HC is capable of producing high quality compressed images. The HC produces about 3 dB improvement in terms of PSNR over standard JPEG. Sections II and III below introduce AVQ, DCT-based TC, and MR-ICS-JPEG. The structure of the new hybrid coder is discussed in section IV. Finally, simulation results and conclusions are given in sections V and VI, respectively.

## Ⅱ. Adaptive Vector Quantization

Because the human eye does not respond with the same sensitivity to low and high detail regions, it is not necessary to allocate the same bit rate to regions having different activities. Therefore, low detail regions are more coarsely coded by adaptive vector quantization (AVQ). AVQ is described below.

Instead of classifying image blocks into edge classes, each block is classified by its homogeneity and coded by a specific codebook which is based on homogeneity. In general, a homogeneous block or region can be thought of as having only a small amount of high frequency components. If we examine the variation of *spatial frequency (SPF)* after a fixed amount of high frequency components is removed, the homogeneity of a given block or region can be measured. In other words, if a low *SPF* is achieved after removing a fixed amount of high frequency components, we can conclude that the given area is homogeneous. The spatial frequency (SPF) of a block is defined as in Eq. 3.

*row frequency* =

$$\frac{1}{N^2}\sum_{i=0}^{N-2}\sum_{j=0}^{N-1}|x_{i,j}-x_{i+1,j}|. \qquad (1)$$

*column frequency* =

$$\frac{1}{N^2}\sum_{i=0}^{N-1}\sum_{j=0}^{N-2}|x_{i,j}-x_{i,j+1}|. \qquad (2)$$

*spatial frequency* =

$$\sqrt{(row\ freq)^2+(col\ freq)^2}. \qquad (3)$$

$x_{i,j}$ : each pixel in spatial domain.

In order to remove high frequency components from a block, the energy compaction property of the DCT is used. After computing the DCT of a block, 25% of the number of coefficients in the high frequency area are removed as illustrated in Figure 1. The shaded areas of Figure 1 represent the removal of high frequency components. The homogeneity based on SPF is defined in Eq. 4.

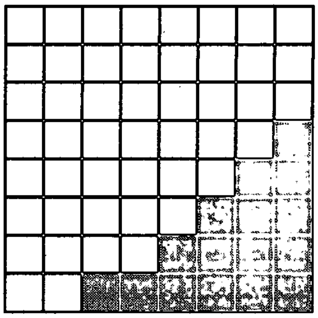


**Fig. 1.** Removal of high frequency components.

$$HSP=\frac{SPFO-SPFR}{SPFO}\times 100\ (\%) \qquad (4)$$

(when $SPFO$ $SPFR$ )

where $HSP$ : homogeneity based on spatial frequency,
$SPFO$ : spatial frequency of original block, and
$SPFR$ : spatial frequency of high frequency removed block.

HSP can be used to describe the activity of each block. The value of SPFR is always less than or equal to SPFO. In Eq. 4, the value of SPFR must be low if a block has few high frequency components. As a result, the homogeneity, HSP, should be high. If coefficients in the high frequency area are originally zero before the removal, SPF does not change after removal. In this case, the homogeneity becomes 100% with the constraint of 25% removal. A set of homogeneous codebooks was generated from sixty-four 256×256 images (65,536 training blocks) of different characteristics using the classical *LBG* algorithm[7]. Training vectors were classified into seven classes according to their homogeneity, and each codebook was generated by the homogeneous training vectors. Table 1 and Figure 2 show the classification and the homogeneity distribution of the training vectors, respectively. Figure 3 represents the PSNR when the training vectors are coded by their corresponding homogeneous codebooks. As can be seen from the graph, the performance is good for highly homogeneous training vectors. Figure 4 presents the block diagram of the adaptive VQ scheme. Adaptive VQ is applied to only low detail (highly homogeneous) blocks in the hybrid coder.

## Ⅲ. Mean Removed Important Coefficient Selection-JPEG

One of the most important factors affecting the performance of an image compression scheme is how well the quantization table

**Table 1.** Classification of training vectors.

| Training Set | Homogeneity Interval (HI) | Average Homogeneity (%) |
|---|---|---|
| Set 1 | 10<HI <=15 | 12.9107 |
| Set 2 | 15<HI<=20 | 17.9660 |
| Set 3 | 20<HI<=25 | 22.8752 |
| Set 4 | 25<HI<=30 | 27.8299 |
| Set 5 | 30<HI<=35 | 32.7365 |
| Set 6 | 35<HI<=40 | 37.7343 |
| Set 7 | HI>40 | 43.2458 |

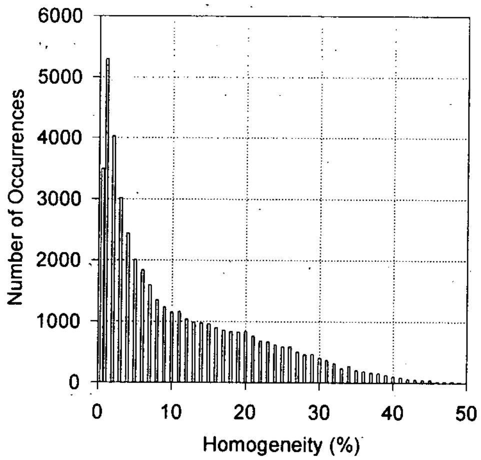


**Fig. 2.** Homogeneity distribution of the training blocks.

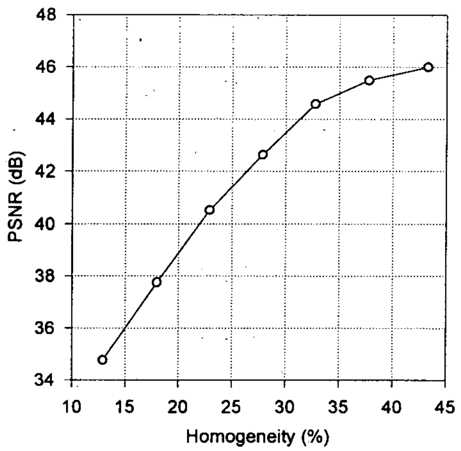


**Fig. 3.** Performance of homogeneous codebooks (8×8 block (training vectors)).

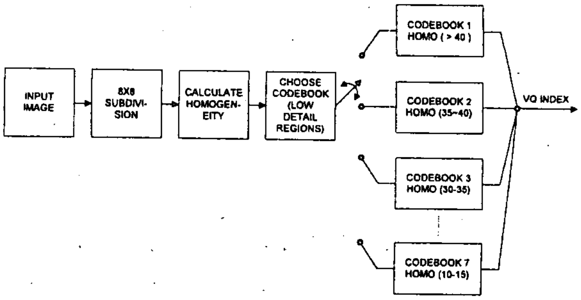


**Fig. 4.** Block diagram of adaptive VQ (AVQ).

matches the data being compressed[8, 9]. This section describes a new method, *Important Coefficient Selection (ICS),* for constructing a quantization table which is shown empirically to give better performance based on maximizing the PSNR. PSNR is defined as follows:

$$PSNR = 10 \log_{10}\left(\frac{255^2}{MSE}\right) \tag{5}$$

The construction technique makes use of the orthonormal property of the DCT for maximizing the PSNR. In the following paragraphs we present this technique, along with several variations, followed by results obtained experimenatlly with JPEG.

### *1. Important Coefficient Selection(ICS) Method*

The Important Coefficient Selection (ICS) method exploits the fundamental role of each DCT coefficient in an 8x8 standard size block for maximizing the PSNR. In the DCT domain, the values of the AC coefficients represent the interpixel activities, and the values of DC coefficients represent the block size times the corresponding block means. The basic idea of the ICS method is based on the orthonormal property of DCT.

The orthonormal property of the DCT results in the same total energy in both data and coefficient domains[2, 4]. This property is expressed in Eq. 6.

$$\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}|x_{ij}|^2 = \sum_{u=0}^{N-1}\sum_{v=0}^{N-1}|X_{uv}|^2 \tag{6}$$

where $x_{ij}$ : picture elements in data domain
$X_{uv}$ : coefficients in DCT domain

According to Eq. 6, the resulting square error is exactly $|X_{uv}|^2$ when we remove one of the $X_{uv}$. Therefore, the squared magnitude of each DCT coefficient represents its generalized degree of importance in a block. The highest squared magnitude in a DCT block corresponds to the DCT coefficient that is the most important in a given block, while the lowest squared magnitude corresponds to the DCT coefficient which is the least important in terms of preserving PSNR. The highest squared magnitude in a DCT block is the coefficient that, if removed, would impact the PSNR most significantly.

We can develop, based on the squared magnitude of DCT coefficients, a new matrix which represents the ranking of impor-

**Fig. 5.** Block diagram for generating ICS-QT.

**Fig. 6.** ICS-Quantization table.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| N/A | 10 | 16 | 20 | 23 | 24 | 25 | 30 |
| 11 | 17 | 21 | 24 | 26 | 27 | 28 | 32 |
| 18 | 21 | 24 | 27 | 28 | 29 | 29 | 33 |
| 23 | 27 | 30 | 31 | 32 | 32 | 31 | 35 |
| 29 | 32 | 34 | 36 | 35 | 35 | 34 | 37 |
| 34 | 37 | 39 | 40 | 39 | 37 | 35 | 38 |
| 38 | 41 | 42 | 41 | 39 | 37 | 39 | 39 |
| 41 | 44 | 44 | 44 | 42 | 40 | 38 | 39 |

(N/A : Not Applicable : DC value is zero, block mean is transmitted in spatial domain)

tance of each DCT coefficient in a block. The DCT coefficients corresponding to high ranking (high squared magnitude) should be retained more accurately after quantization while those having low ranking become candidates to be quantized coarsely or removed. The ranking procedure is performed by sorting the squared DCT coefficients of a block into descending order. The coefficient which has the highest squared magnitude in a DCT block is assigned the ranking 1 while the coefficient having the lowest squared magnitude corresponds to ranking 63. The ranking matrix results in a quantization table constructed so as to maximize block PSNR.

In order to generate the quantization table for our tests, we used a set of seven $576 \times 720$ gray scale images (a total of 45,360 $8 \times 8$ blocks) for training. The arithmetic mean was used to merge the 45,360 8x8 ranking matrixes for constructing the quantization table. Figures 5 and 6 show a block diagram which illustrates the ICS-QT method and the corresponding ICS quantization, respectively.

### 2. *Mean Removed Quantization*

In the DCT transformed domain, a given DC coefficient has the value of block size times the corresponding block mean. Therefore, the possible magnitudes of DC coefficients range from 0 to 2,048 for an 8-bit image when the block size is $8 \times 8$. As a result, even though DPCM[10] may be used to code a DC coefficient, sometimes 8 bits or even 9 bits may not be enough to represent accurately the DC value.

If the mean of each block is removed, the statistical expected value of each block is zero and the maximum block mean in the *spatial domain* is less than or equal to 255 for 8-bit images which is much smaller than the magnitude of the average DC coefficient in the transform domain. Therefore we can achieve more compression by removing block means; the values of the DC coefficients will all be zero while the activities among adjacent AC coefficients in the DCT domain will remain the same. The DCT transformed residual vectors are quantized using the quantization table which was generated using the ICS method described in the previous section. The block means are transmitted to the decoder using DPCM in the *spatial domain.* Finally, the appropriate mean values are added to the corresponding reconstructed residual vectors. The mean removal procedure and the ICS-QT are directly applied to the JPEG baseline algorithm. Therefore, it is referred to as MR-ICS-JPEG here. Figure 7 illustrates, by means of a block diagram, the MR-ICS-JPEG procedure.

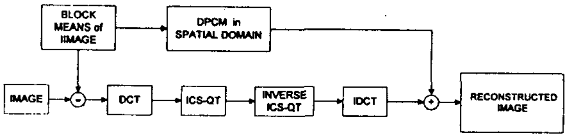


**Fig. 7.** Block diagram of MR-ICS-JPEG.

## Ⅳ. Structure of a New Hybrid Coder

This section presents the structure of the new hybrid coder which combines the AVQ and MR-ICS-JPEG. The high overhead and design complexity of edge classifiers have been limitations of other published hybrid coders[3, 4, 5, 6]. The new hybrid coder provides low overhead and a simple classifier using homogeneity. In addition, the problems of VQ and DCT based transform coder, encoding complexity and low compression, are alleviated by combining these two complementary compression schemes.

Table 2 represents the possible candidates for hybridization. Because the best performance of DCT based transform coding is achieved at block size $8 \times 8$[9], methods II and IV are not good choices. Moreover, method IV requires too many overhead bits which are produced by quadtree segmentation for handling different sizes of blocks and adaptive VQ. Therefore, the combination of quadtree and adaptive schemes is not a good choice. Method III produces reasonable overhead; however, AVQ requires too many bits to code blocks as small as $2 \times 2$. In method V, the error vectors are usually more random than the original ones. As a result, it is hard to achieve energy compaction from the DCT coding step. Therefore, the objective qualities from methods II to V in terms of PSNR are lower than that of MR-ICS-JPEG alone.

On the other hand, Method I produces no more than 3 bits of overhead, and more bits can be allocated to MR-ICS-JPEG because low detail 8x8 blocks are coded by AVQ. As a result, Method I produces better performance than when MR-ICS-JPEG is used alone. Figure 8 shows the simulation result of Method I. As can be seen from Figure 8, Method I produces about a 3.5 dB performance improvement compared to that of standard JPEG.

**Table 2.** Possible candidate for Hybridization.

| Block Size \ Method | | I | II | III | IV | V |
|---|---|---|---|---|---|---|
| 8×8 | Low Detail | AVQ | AVQ | MR-ICS-JPEG | AVQ | Multistage coding:<br>1st stage: HVQ<br>2nd stage: MR-ICS-JPEG |
| | High Detail | MR-ICS-JPEG | | | | |
| 4 × 4 | | x | MR-ICS-JPEG | x | MR-ICS-JPEG | |
| 2 × 2 | | x | x | AVQ | AVQ | |

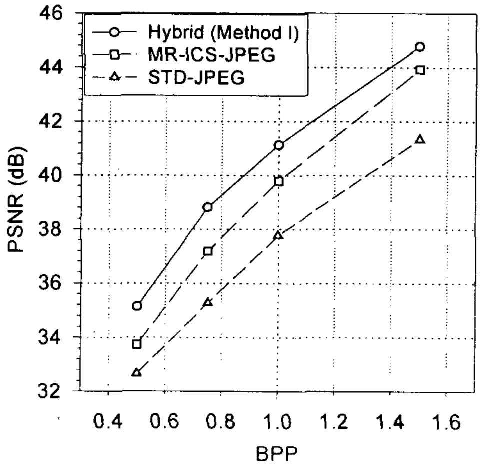


**Fig. 8.** Performance of Method I and other compression schemes (Lenna 256×256).

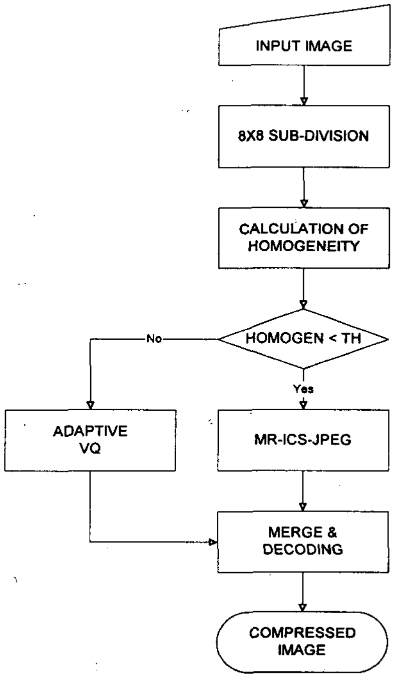


**Fig. 9.** Flow chart of the new hybrid coder.

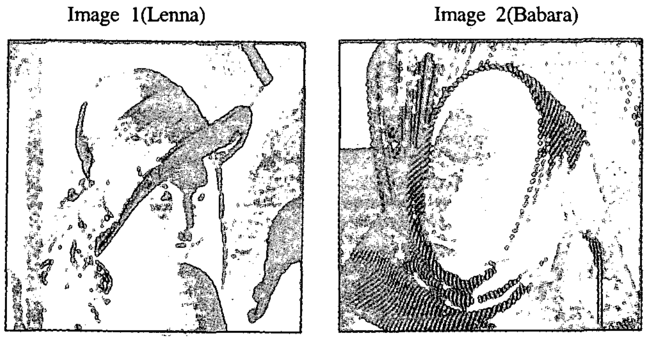
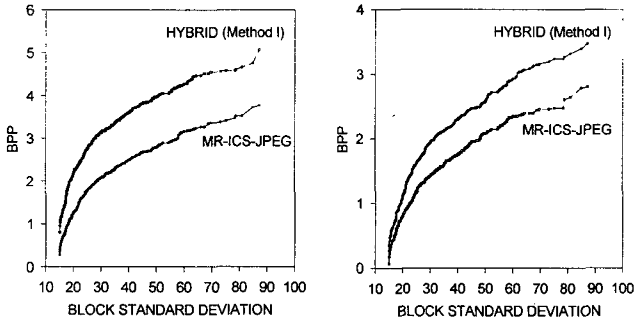
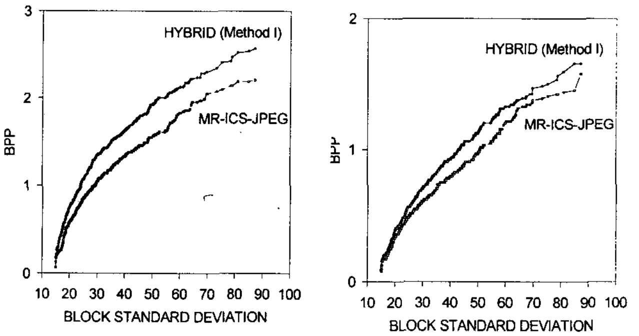


**Fig. 10.** Test images (Lenna and Barbara).



(a) 1.5 bpp (b) 1.0 bpp



(c) 0.75 bpp (d) 0.5 bpp

**Fig. 11.** Allocation of additional bits to *MR-ICS-JPEG*(Lenna).

Figure 9 presents the flow chart of the new hybrid coder.

## V. Discussion of Simulation Results

Simulation is performed for two test images, Lenna and Barbara, which are shown in Figure 10. Lenna is included in the training vectors while Barbara is outside the training vectors. One of the advantages of the hybrid coder is its adaptivity. In other words, low detail regions are coded roughly by AVQ while high detail regions are coded finely by MR-ICS-JPEG to represent the edges or curves more accurately. By this strategy, the hybrid coder produces not only objectively, but also subjectively, enhanced quality of compressed images. Figure 11 represents how many additional bits can be allocated to high detail regions using the

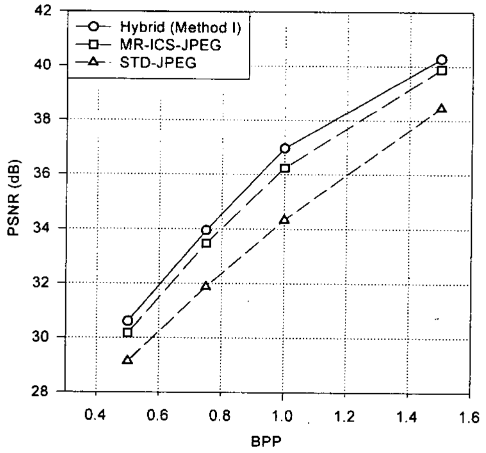
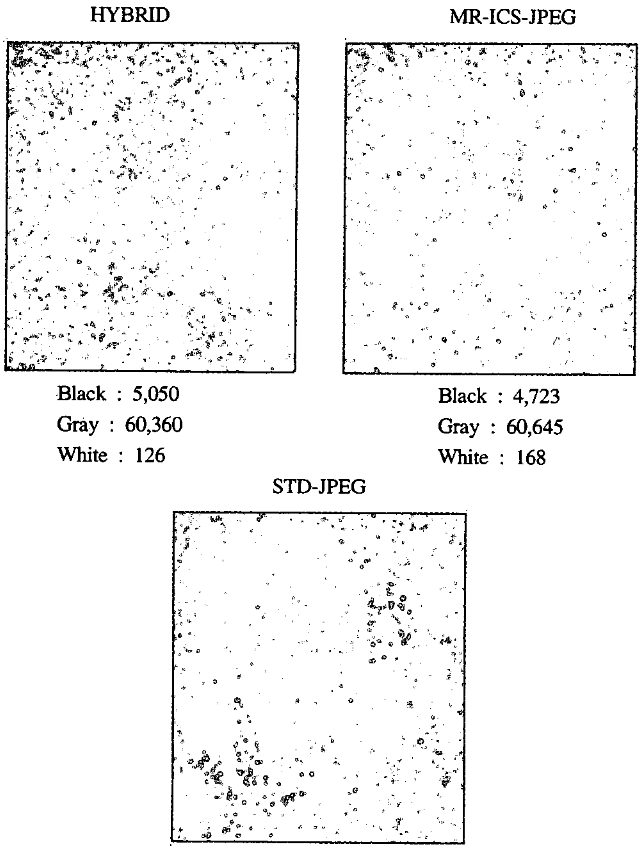


**Fig. 12.** Performances of different compression schemes(Barbara).



HYBRID

Black : 5,050
Gray : 60,360
White : 126

MR-ICS-JPEG

Black : 4,723
Gray : 60,645
White : 168

STD-JPEG

Black : 4,676, Gray : 60,261, White : 599

**Fig. 13.** Error image of Barbara (0.5bpp) (zero error (black), less than or equals to 30 (gray), and larger than 30 (white)).

hybrid scheme compared to when MR-ICS-JPEG is used alone. The x-axis represents the sequence of blocks, which are coded by MR-ICS-JPEG and hybrid coder, sorted in terms of block standard deviation. As can be seen in Figure 11, approximately 1.5, 1.0, 0.5, and 0.25 bpp can additionally be allocated to *MR-ICS-JPEG* at the bit rates of 1.5, 1.0, 0.75, and 0.5 bpp, respectively.

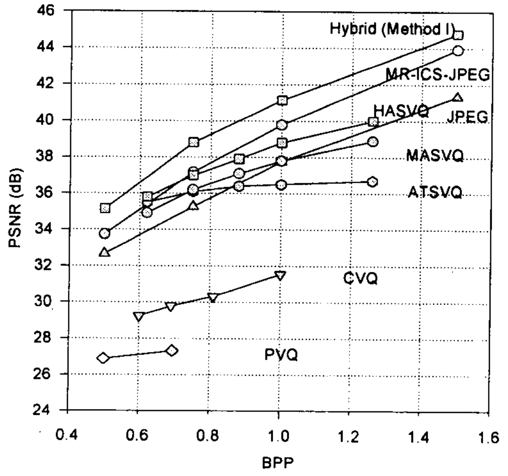


**Fig. 14.** Comparison of several comparable compression schemes (Lenna) (HASVQ (Hierarchical Adaptive Search VQ), MASVQ (Multi-Stage Adaptive Search VQ), and ATSVQ (Adaptive Tree Search VQ) [11]. CVQ (Classified VQ) [12]. PVQ (Pyramid VQ) [13].)

This is the main source for enhancement of the objective and subjective qualities.

When an input image is outside the training vectors, the performance in terms of PSNR of the hybrid coder is lower than when an input image is inside the training vectors. Figure 12 shows the performances in terms of PSNR of the Barbara image which is outside the training vectors. The performance of the hybrid coder is about 2.5 dB higher than that of JPEG. Further, informal subjective testing produced noticeable improvement with the hybrid coder when compared to JPEG alone. This is due to additional bits which were made available by AVQ which could then be allocated to high detail active regions.

Figure 13 represents the error images of Barbara which were generated by different schemes. In Figure 13, pixels which produce zero error, less than or equal to 30, and larger than 30 are displayed by pure black, gray, and white, respectively. In this comparison, the hybrid coder produces the fewest white spots which represent pixel error larger than 30.

The other advantage of the proposed hybrid coder is the low overhead in bits. Maximum overhead in an 8x8 block is restricted to 3 bits in this scheme. In order to verify the superiority of the proposed schemes, hybrid coder and MR-ICS-JPEG, direct comparisons in terms of PSNR with other published results are given in Figure 14. As can be seen from Figure 14, the hybrid coder produces the best performance in terms of PSNR, followed closely by MR-ICS-JPEG.

## VI. Conclusions

A new hybrid coder which combines MR-ICS-JPEG and AVQ was described and simulated. This coder was shown empirically to produce about 3.5 dB and 2.5 dB improved performances when an input image is inside and outside of the training vectors, respectively, compared to JPEG. Although image blocks are not classified by their edge patterns, the codebook adaptive VQ scheme represents low detail regions appropriately.

The advantages of the proposed hybrid coder are its adaptivity and low overhead in bits. In other words, low detail regions are coded roughly by AVQ, and some additional bits, which are made available by AVQ, are allocated to the MR-ICS-JPEG stage to represent high detail regions more accurately. This strategy and low overhead produce subjectively and objectively enhanced image quality.

## References

[1] A. Gersho and R. M. Gray, Vector Quantization and Signal Compression, Kluwer Academic Publishers, Boston, 1991.

[2] R. J. Clarke, Transform Coding of Images, Page 4, Academic Press, London, 1985.

[3] K. N. Ngan, H. L. Koh, and W. C. Wong, Hybrid Image Coding Scheme incorporating Human Visual System Characteristics, Optical Engineering, Vol. 30, No.7, pp. 940-946, July 1991.

[4] K. A. Wen and C. Y. Lu, Hybrid Vector Quantization, Optical Engineering, Vol. 32, No. 7, pp. 1496-1502, July 1993.

[5] M. Ohta and S. Nogaki, Hybrid Picture Coding with Wavelet Transform and overlapped Motion-Compensated Interframe Prediction Coding, IEEE Transactions on Signal Processing, Vol. 41, No. 12, pp. 3416-3424, Dec. 1993.

[6] M. J. Bage, Interframe Predictive Coding of Images Using Hybrid Vector Quantization, IEEE Transactions on Communications, Vol. COM-34, No. 4, pp. 411-415, April 1986.

[7] Y. Linde, A. Buzo, and R. M. Gray, An Algorithm for Vector Quantizer Design, IEEE Transactions on Communications, Vol. COM.-28, No.1, pp. 84-95, Jan. 1980.

[8] G. K. Wallace, The JPEG Still Picture Compression Standard, IEEE Transactions on Consumer Electronics. Vol. 38, No. 1, pp. xviii-xxxiv, Feb. 1992.

[9] W. B. Pennebaker and J. L. Mitchell, JPEG Still Image Compression Standard, Van Nostrand, New york, 1993.

[10] A. K. Jain, Fundamentals of Digital Image Compression, pp. 483-495, Prentice-Hall, 1989.

[11] M. A. Ghafourian and C. H. Huang, Comparison between several Adaptive Search Vector Quantization Schemes and Standard for Image Compression, IEEE Transactions on Communications, Vol. 43, No. 2/3/4, pp. 1308-1312, Feb./Mar./Apr. 1995.

[12] B. Ramamurth and A. Gersho, Classified Vector Quantization of Images, IEEE Transactions on Communications, Vol. Com-34, No. 11, pp. 1105-1115, Nov. 1986.

[13] J. W. Kim and S. U. Lee, A Transform Domain Classified Vector quantizer for Image Coding, IEEE Transactions on Circuits and Systems for Video Technology, Vol. 2, No. 1, pp. 3-14, Mar. 1992.

**Hang-Chan Lee** was received the B.S. degree in electrical engineering from Inha University, Incheon, Korea in 1986, and received the M.S. and Ph.D. in electrical and Computer engineering from Oklahoma State University, U.S.A., in 1992 and 1997, respectively. He is currently a member of research staff at SAIT. His research interests include digital image processing, computer vision, and digital communications.