[25] Y. Linde, A. Buzo, and R. M. Gray, "An algorithm for vector quantizer design," *IEEE Trans. Commun.*, vol. C-28, pp. 84–95, Jan. 1980.

# Improvement on Image Transform Coding by Reducing Interblock Correlation

Kyoung Won Lim, Kang Wook Chun, and Jong Beom Ra

***Abstract*—In this correspondence, we try to improve transform coding efficiency by alleviating interblock correlation due to the small size of block. The proposed method needs minor modification from conventional transform coding techniques such as JPEG, and reduces information loss in the coding procedure for a given bit rate. Simulation results demonstrate that the method drastically diminishes the blocking effects and enhances the subjective visual quality compared with such existing algorithms as JPEG and LOT.**

## I. Introduction

In still image coding applications such as color facsimile, video text, photo telegraphy, etc., the block transform coding technique using discrete cosine transform (DCT) is extensively used. For transform image coding [1]–[3], an image is divided into many subimages or blocks before coding, and each equal size block is coded independently. By partitioning the image, the coder can adapt to local image characteristics and its hardware implementation is simplified. In sufficiently low bit rates, however, transform coding introduces typical image degradation, i.e., blocking effects between subimages, which is known as its major drawback [4]–[6]. The visibility of these effects depends on local image characteristics and these effects are specially notable in the smooth region.

Several methods to reduce the blocking effects have been suggested previously [2], [4]–[8]. Among them, the prediction method replaces some low-frequency DCT coefficients with the DCT coefficients predicted from the second-order polynomial surface model using DC values of nine subimage blocks, including the eight nearest-neighbor blocks at the receiver [2]. Although the method does not increase the bit rate, it distorts the image edges inherently. As another approach to reduce the blocking effects, the lapped orthogonal transform (LOT) was proposed [6], [7]. Since the LOT of a block depends on the neighboring blocks, however, it introduces more ringing artifacts around the edges, and adds the computational burden. Recently, the combined-transform coding scheme was also suggested [8]. The algorithm combines noiseless coding for the upper (bit) image set without dividing into blocks and conventional block transform coding for the lower image set. It is noted that its coding efficiency is closely related to the correlation of upper image set.

In this correspondence, we propose the improved DCT transform coding techniques that retain the advantages of block-based coding and diminish the visibility of blocking effects by exploiting and alleviating interblock correlation. To initiate the problem, the conventional transform coder, JPEG, is briefly discussed in Section II. In Section III, the proposed algorithm is described. The performance of this scheme is examined with a test image and compared with those of other coding algorithms in Section IV.

Manuscript received July 7, 1993; revised October 18, 1994. This work was supported by MunWha Broadcasting Corporation (MBC), Korea. The associate editor coordinating the review of this paper and approving it for publication was Prof. William A. Pearlman.

The authors are with the Department of Electrical Engineering, Korea Advanced Institute of Technology, 373-1, Kusongdong, Yusonggu, Taejon, Korea.

IEEE Log Number 9412459.

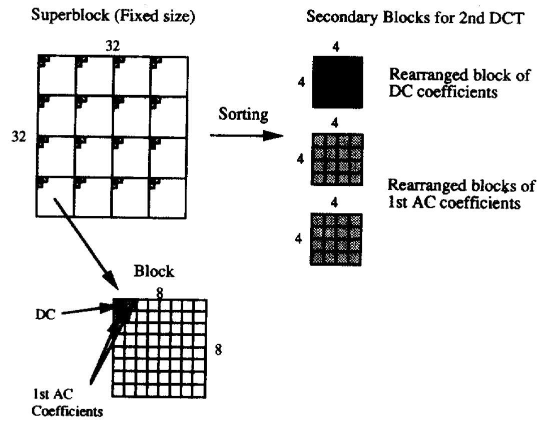



Fig. 1. Fixed-size superblock structure after 2-D $8 \times 8$ DCT.

## II. The Conventional Transform Coder

The JPEG baseline system is a typical example of the conventional transform coder. The system consists of three major parts, namely, transformation, coefficient quantization, and variable length encoding.

The transformation converts the space-domain block to the frequency-domain one. The DCT is used for transformation because of its excellent energy compaction and computation efficiency. The block size for DCT is usually chosen as $8 \times 8$ pixels to compromise between its computational complexity and energy compaction performance. The DCT coefficients in the block are then quantized using $8 \times 8$ matrix quantization values that have been determined by observing the sensitivity of the human vision to the quantization errors. The quantized output is fed to a zigzag scanning procedure, and converted into a set of symbols. Then, the symbols are encoded by using the variable-length Huffman coder. In the decoder, the subimage block is reconstructed in the reverse order. Since each subimage block is coded independently in the JPEG baseline system, the objectionable blocking effects are introduced, and in turn, these effects cause major image degradation as the bit rate decreases.

## III. The Proposed Transform Coder

Due to the small block size in transform coding, the global characteristics of the image cannot be reflected in each block transformation. As an example, blocks in the smooth region of the image consist of highly correlated pixels, and the energy of each block is concentrated on a few lower frequency DCT coefficients. Since the interblock correlation still remains high, however, the DCT coefficients at the same frequency in the blocks are expected to have strong correlation with each other. Since the visibility of blocking effects in the smooth region is higher than in the complex region such as edge and texture, if the interblock correlation is utilized for more effective coding, the visual quality of the reconstructed image can be improved.

To examine the interblock correlation of the DCT coefficients, the superblock, or a set of blocks in proximity, is defined in the DCT domain as shown in Fig. 1. The superblock has a size of $4 \times 4$ DCT blocks (each block has $8 \times 8$ pixels) and the DCT coefficients in the same locations are picked up from each block in a superblock and rearranged into the secondary blocks. To examine the interblock

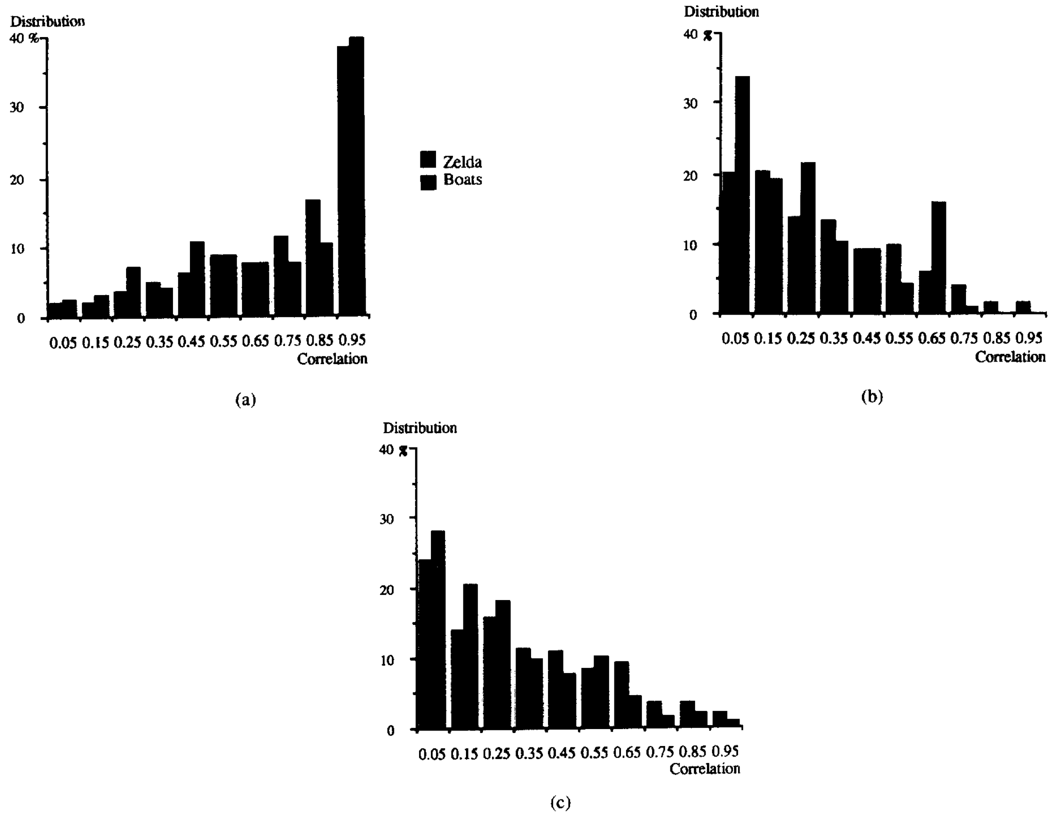



Fig. 2. Correlation value distributions of (a) dc coefficient; (b) first ac coefficient; $A_{01}$; (c) first ac coefficient, $A_{10}$.

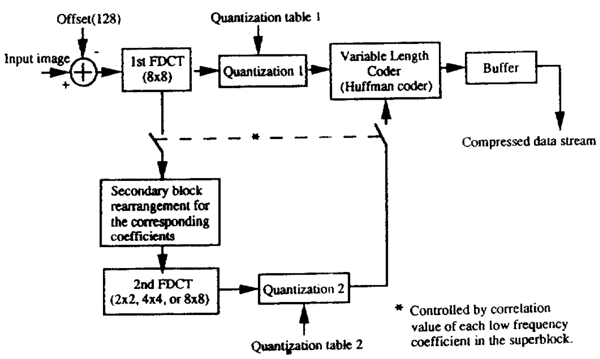



Fig. 3. Encoder block diagram of the proposed method.

correlation, the correlation value $\rho$ is measured in the secondary block as follows:

$$\rho = \min\left[\left|\frac{r(0,1)}{r(0,0)}\right|, \left|\frac{r(1,0)}{r(0,0)}\right|\right] \tag{1}$$

where

$$r(k,l) = \frac{\sum_{m=0}^{N-1-k}\sum_{n=0}^{N-1-l} F'(m,n)F'(m+k,n+l)}{(N-k)(N-l)} \tag{2}$$

$$F'(m,n) = F(m,n) - \mu \tag{3}$$

and

$$\mu = \frac{1}{N^2}\sum_{m=0}^{N-1}\sum_{n=0}^{N-1} F(m,n). \tag{4}$$

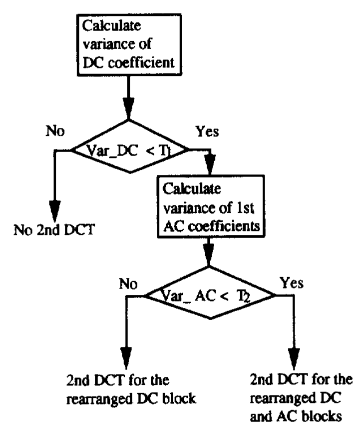



Fig. 4. Second DCT procedure for the fixed-size superblock. The superblock structure is given in Fig. 1.

Here $F(m,n)$, $\mu$, and $r(\cdot)$ are the element, mean, and autocovariance of an $N \times N$ secondary block, respectively, and $N = 4$ in Fig. 1. Fig. 2 shows the correlation values of dc and the first ac coefficients in some images, which are sufficiently high (more than 0.6) in 70~80% and 15~20% of superblocks, respectively. These results suggest that the selective removal of interblock correlations of the dc and first ac coefficients can improve the transform coding efficiency.

Fig. 3 shows the overall coding scheme of the proposed transform coder. The first $8 \times 8$ DCT is performed on the input subimages, and

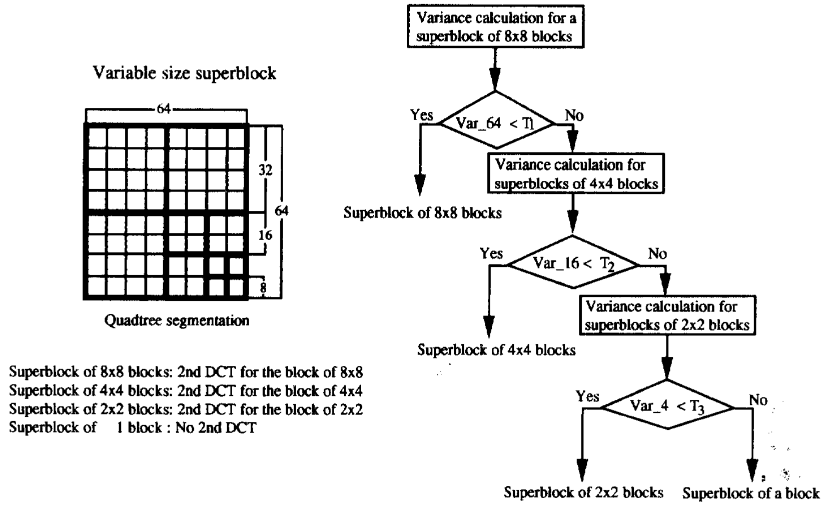



Fig. 5. Superblock size determination in the variable block case.

Luminance

| | | | |
|---|---|---|---|
| 8 | 10 | 24 | 51 |
| 13 | 24 | 57 | 56 |
| 18 | 37 | 68 | 103 |
| 64 | 87 | 121 | 101 |

Chrominance

| | | | |
|---|---|---|---|
| 17 | 24 | 66 | 99 |
| 26 | 99 | 99 | 99 |
| 99 | 99 | 99 | 99 |
| 99 | 99 | 99 | 99 |

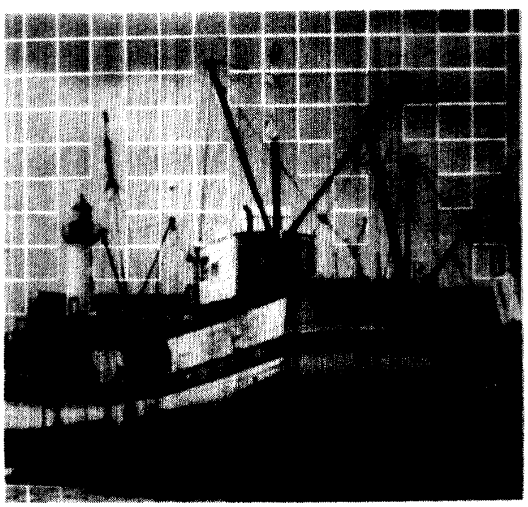
(a)

Luminance and Chrominance

| | |
|---|---|
| 8 | 6 |
| 6 | 12 |

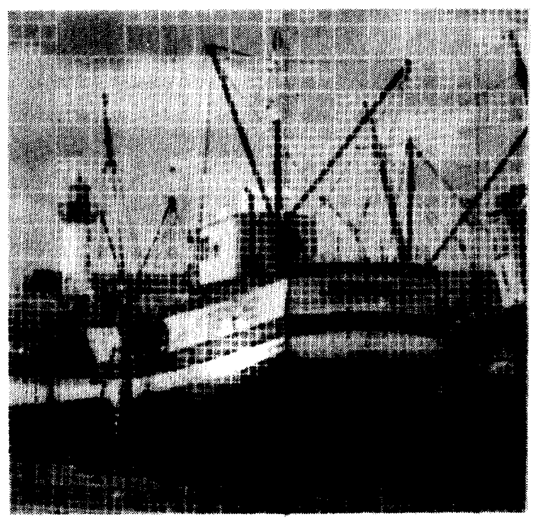
(b)

Fig. 6. Quantization tables for the second DCT coefficients. (The one for $8 \times 8$ pixels is identical with the JPEG default quantization table.) (a) $4 \times 4$. (b) $2 \times 2$.



(a)

Fig. 7. Superblocks classified for the second DCT. (a) Fixed size case of $4 \times 4$ blocks. (b) Variable size case.

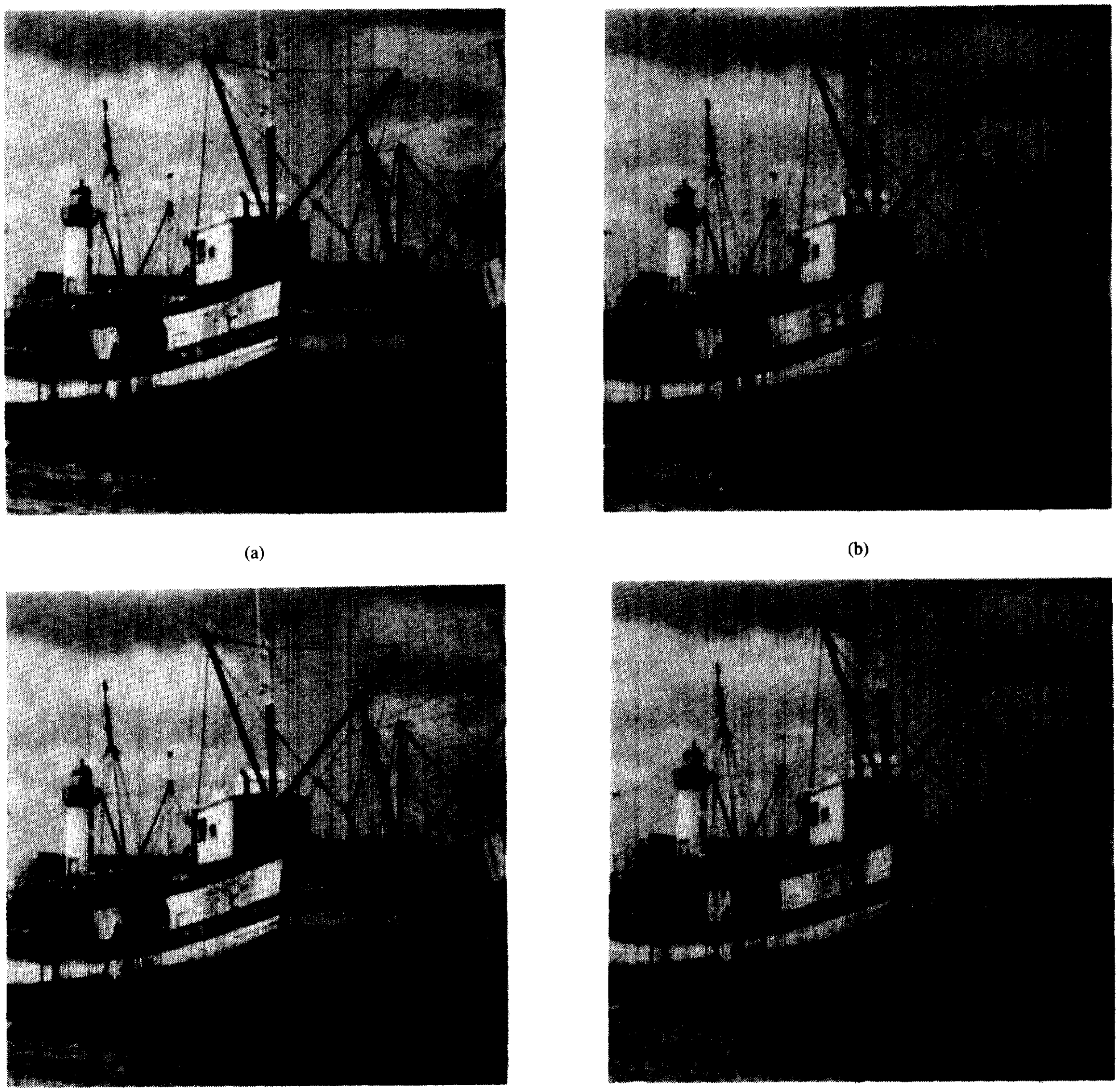


(b)

Fig. 7. *(Continued)*

then superblocks are formed in the DCT domain (refer to Fig. 1). Interblock correlations for the dc and two first ac coefficients in each superblock are examined. If the correlation is high enough, the corresponding coefficients are rearranged onto a new block and the second DCT is performed. It should be noted that the block for the second DCT consists of DCT coefficients of the neighboring blocks at the same DCT-frequency location from the first $8 \times 8$ DCT. The coefficients remaining in the first transformed domain and the coefficients in the second transformed domain are coded according to the algorithm based on JPEG, respectively.

The decision to use interblock coding for the superblock coefficients is very important to the coding performance. Since a highly correlated superblock has low variance, the variance of each component is adopted as the decision criterion. To exploit the interblock correlation, two types of superblock, i.e., the fixed-size and variable-size superblocks, have been examined with different procedures. For the fixed superblock size (see Fig. 1), if the dc coefficient has low variance, the variances of the other two coefficients are successively examined, and interblock coding is performed for the coefficients of



(a) (b) (c) (d)

Fig. 8. Reconstructed images at 0.48 bpp. (a) JPEG. (b) LOT. (c) Proposed method with the superblock of fixed size. (d) Proposed method with the superblock of variable size (the quadtree segmentation overhead is 1% of the bit rate).

low variance (see Fig. 4). If the superblock is constructed in the form of quadtree and its size when investigating the interblock correlation is adjusted according to scene complexity, however, we can use the interblock correlation for coding more effectively. Fig. 5 shows the interblock coding procedure for the variable size superblock. The interblock correlation values of the dc and two first ac coefficients are examined for the superblock of $8 \times 8$ blocks (or $64 \times 64$ pixels) first, and, if they are high enough, interblock coding is performed. Otherwise, the $8 \times 8$ blocks are divided into four smaller superblocks of $4 \times 4$ blocks, and the same processes are repeated.

The quantization table 1 in Fig. 3 is adopted from JPEG [2]. The quantization table 2 has three sets depending on the secondary block size. For the $8 \times 8$ block, the quantization table 1 is used without modification, and the tables for the $4 \times 4$ and $2 \times 2$ blocks are determined based on the subjective quality test (see Fig. 6).

## IV. Experimental Results

To verify the performance of the proposed algorithm, the "Boats" image, which conforms to CCIR 601 format, is used. The image has various types of scene complexity such as smooth region, edge, texture, etc. Fig. 7(a) shows the superblocks classified for the second DCT in the case of the fixed size superblock of $4 \times 4$ blocks. Even though this method is simple to implement, the coding efficiency is reduced if the superblock includes abrupt local changes. The variable superblock size can alleviate this problem, and its classification result is shown in Fig. 7(b).

Performance evaluation has been carried out for the original JPEG, its blocking effect reduction method, LOT, and the proposed methods with a bit rate of 0.48 bpp. It is noted in Table I that PSNR's of the proposed methods are not higher than those of other methods.

TABLE I
PSNR'S AT THE BIT RATE OF 0.48 BPP.

| | JPEG | JPEG* | LOT | Method1 | Method2 |
|---|---|---|---|---|---|
| PSNR(dB) | 34.2 | 34.0 | 34.4 | 34.2 | 34.2 |

JPEG* : AC prediction method in JPEG
Method 1: Proposed method with fixed size superblock
Method 2: Proposed method with variable size superblock

However, blocking effects in the smooth area are reduced and the subjective image quality is enhanced substantially, as shown in Fig. 8. Even though the second DCT and the correlation value measurement are to be additionally performed on several coefficients in the transform domain, the computational overhead is at most 9% more than that of the original JPEG.

## V. CONCLUSION

We proposed the improved transform coding technique by reducing the interblock correlation, and also evaluated its performance in comparison with other existing algorithms. The proposed method provides the substantial visual quality improvement in the reconstructed image for a given bit rate. Since the method requires just minor modification from the conventional transform coder, it seems a very promising transform coding method at a low bit rate.

## REFERENCES

[1] P. A. Winz, "Transform picture coding," *Proc. IEEE*, vol. PROC-60, pp. 809–820, July 1972.
[2] *JPEG Technical Specification, Revision 5*, JPEG-8-R5, Jan. 1990.
[3] J. S. Lim, *Two-Dimensional Signal and Image Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1990.
[4] H. C. Reeve, III and J. S. Lim, "Reduction of blocking effect in image coding," in *Proc. ICASSP '83*, Boston, MA, 1983, pp. 1212-1215.
[5] P. M. Farrelle and A. K. Jain, "Recursive block coding—A new approach to transform coding," *IEEE Trans. Commun.*, vol. COM-34, pp. 161–179, Feb. 1986.
[6] H. S. Malvar and D. H. Staelin, "The LOT: Transform coding without blocking effect," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, pp. 553–559, Apr. 1989.
[7] P. M. Cassereau, D. H. Staelin, and G. Jager, "Encoding of images based on a lapped orthogonal transform," *IEEE Trans. Commun.*, vol. 37, pp. 189–193, Feb. 1989.
[8] Y. Q. Zhang, R. L. Pickholtz, and M. H. Loew, "A new approach to reduce the blocking effect of transform coding," *IEEE Trans. Commun.*, vol. 41, no. 2, Feb. 1993.

# Low Entropy Image Pyramids for Efficient Lossless Coding

David Houlding and Jacques Vaisey

***Abstract*—An efficient image source coding technique gives good compression performance at low computational complexity. This research introduces an efficient coding technique, based on pyramid coding, that involves transforming an image into an equivalent lower entropy form prior to lossless coding. The proposed method is also a multiresolution technique that facilitates progressive image transmission.**

## I. INTRODUCTION

The pyramid generation process permits the use of a large variety of decimation and interpolation filters. The choice of these filters has a significant influence on the properties of the generated pyramid [1]. In order to achieve good lossless compression, it is desirable to generate a pyramid in which the subimages have low entropy.

In the decimation process, an image is first lowpass filtered by a 2–D decimation filter, after which it is subsampled by a factor of two horizontally and vertically. The decimated image contains the low spatial frequency information present in the original image but has only one fourth the number of pixels. The purpose of the decimation filter is to reduce aliasing by attenuating spatial frequencies above $\pi/2$ in the original image.

The interpolation process consists of upsampling a decimated image by a factor of two horizontally and vertically, followed by lowpass filtering of the upsampled image using an interpolation filter. The purpose of the interpolation filter is to amplify the spatial frequencies below $\pi/2$ in the upsampled image that are attenuated as a result of the upsampling operation and to attenuate the spatial frequencies above $\pi/2$ representing the unwanted "interpolation image."

The pyramid generation iteration consists of decimation, followed by interpolation and subtraction of the interpolated image from the original. The difference and decimated images form the base and top of a two level pyramid and contain mostly high and low spatial frequency information, respectively. The image reconstruction iteration then consists of interpolating the top of the pyramid and adding it to the base. Larger pyramids can be generated by cascading the generation iteration so that the top of the pyramid formed from the first iteration is the original image for the second iteration and so forth. The reconstruction iteration may be cascaded in a similar manner.

After generation of the pyramid, the subimages are entropy coded. Since it is required that the pyramid generation be lossless, no quantization of the pyramid subimages can occur after generation without the use of more elaborate quantization feedback schemes [2]. However, decimation and interpolation filtering generally produce a wide range of floating point pixel values that make the pyramid subimages unsuitable for entropy coding. In order to ensure that

Manuscript received September 29, 1993; revised September 30, 1994. This work was supported by Canadian Cable Labs and the Natural Sciences and Engineering Research Council of Canada. The associate editor coordinating the review of this paper and approving it for publication was Prof. Nasser M. Nasrabadi.

D. Houlding is with MacDonald Dettweiler, Richmond, BC, Canada.
J. Vaisey is with the School of Engineering Science, Simon Fraser University, Burnaby, BC, Canada, V5A 1S6.

IEEE Log Number 9412464.