Wavelet Transform Based Variable Tree Size Fractal Video Coding

Ying Zhang(1), *Lai-Man Po*(2) and *Ying-Lin Yu*(1)

(1) Dept. of EE, South China University of Technology, Guangzhou, 510641, P. R. China
(2) Department of Electronic Engineering, City University of Hong Kong, Hong Kong

ABSTRACT

A new wavelet video compression scheme using adaptive fractal coding is presented in this paper. Through pyramidal wavelet transform, each video frame is decomposed into multiresolution subbands and then organized into a set of wavelet subtrees to represent motion activities. After multiresolution motion detection, these wavelet subtrees are classified into motion and non-motion subtress. The coding of non-motion subtrees is straightforward and simple, while the motion subtress are adaptively separated and then encoded using variable tree size fractal coding. Experimental results show that the proposed scheme provides a superior performance in terms of PSNR as well as the subjective quality at low bit-rates.

I. INTRODUCTION

Recently, the wavelet and fractal coding has aroused a lot of attention in both still image coding and video compression, and hence become two important branches of image/video coding methods. The wavelet representation provides a multiresolution and multi-frequency expression of nonstationary image signals with localization in both time and frequency domain. Such property is desirable for image and video coding because multiresolution subbands become relatively more stationary and can be coded separately according to local statistical properties [1]. In commonly used wavelet transform video coding scheme [2], the multiresolution frequency subbands are also processed independently so that a high compression ratio is relatively difficult to obtain.

Fractal coding is an alternative new tool to compress the image and video signals. The first automatic blockwise fractal image compression algorithm was presented by Jacquin [3] and has been extensively studied due to its potential high compression ratio and fast decoding performance[4, 5]. All these existing schemes utilize affine transformation which maps one part of an image to another part of the same image in spatial domain. However, the spatial domain based fractal coding scheme suffers from annoying blocking effects especially at low bit-rates which hinders its practical applications.

Nonetheless, how to combine wavelet with fractal coding in one compression scheme is still a new research topic till now. The initial work of generalizing the fractal coding from spatial domain to the wavelet domain was done by G. Davis [6] and H. Krupnik et. al. [7] independently. By constructing wavelet subtree composed of the coefficients at different resolution and orientation but with same spatial location, Davis demonstrated that conventional fractal block codec is a form of Haar wavelet subtree quantization scheme. If smooth wavelet basis is employed, the blocking artifacts are dramatically reduced. In addition, this new approach of wavelet framework leads to develop an unconditionally convergent fractal block coder with a fast decoding algorithm.

In this paper, a new variable wavelet subtree size based fractal coding is developed in section II for still image compression which can obtain a good trade-off between image quality and compression ratio, and then this method is extended in section III to encode video signals which exploits the temporal redundancy among adjacent frames and has achieved good simulation results.

II. VARIABLE WAVELET SUBTREE SIZE BASED FRACTAL CODING

Suppose an image is transformed to the wavelet domain by a 4 stage pyramidal decomposition as shown in Fig. 1, the three coefficients from each of the three high frequency subband at the coarsest scale, which have the same spatial location, together with their children and grandchildren, 2×2, 4×4 and 8×8 coefficients at successive finer scales, are highly correlated and can be formed to a hierarchical data structure — wavelet subtree. The three coefficients at the highest tree layer are called root nodes, and the coefficients at the lowest tree layer are called leaf nodes. The two different size wavelet subtrees with root at scale level 4 and 3 depicted in Fig. 1 are called domain and range subtree, respectively. It is observed that such wavelet subtree possesses a structure of self-similarity which can predict the coefficients of a range subtree from those of a domain subtree [6, 7]. Thus, the main idea of wavelet-based fractal coding is to approximate the range subtree by the best-matched domain subtree through proper affine transformation. In general, the proposed new hybrid coding scheme can be summarized as follows.

To give the initial condition for multiresolution wavelet subtree based prediction, the coefficients of the four

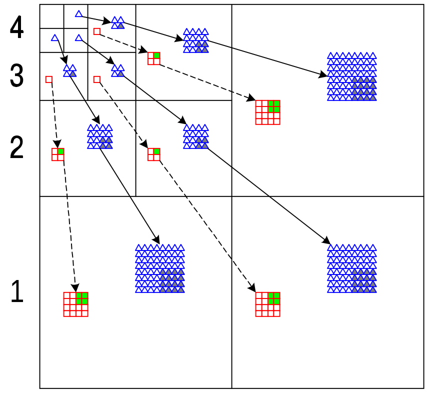

Fig. 1. Construction of wavelet subtrees by a four-stage pyramidal wavelet decomposition. The domain and range subtree consists of the triangular and square pixels, respectively.

subbands at the coarsest scale level of wavelet pyramid are scalar quantized within each subband independently. The rest coefficients are grouped into non-overlapped range subtrees. Meanwhile, the domain subtree with root at coarsest scale level are also constructed to form the searching pool (See Fig. 1). Owing to the special structure of wavelet subtree, the affine transformation used for wavelet subtree mapping is different from that of the conventional spatial domain fractal coding (refer to [6, 7] for details). For a given range subtree, the best-matched domain with the minimum distortion ε is found after full search. To determine whether such approximation is good enough, the distortion ε is compared to a predefined threshold T. If ε is below T, in this case, the range subtree can be "fractal quantized" by the domain subtree within given distortion and further splitting is not needed. On the contrary, if the distortion ε exceeds T, the range subtree is further segmented into children subtrees and then coded separately to reduce the reconstruction error.

As the coefficients at coarser scale usually possess more energy than those at finer scale, the three root nodes are removed from the tree and then scalar quantized for less coding error. The pruned range subtree is split into four quadrant children subtrees, with 3 root nodes at next finer scale level 2. This adaptive partition process is shown in Fig. 1 where the shaded square pixels represent one of the four children subtrees. To code these children subtrees, the corresponding searching pool is constructed consisting of domain subtrees with root nodes at scale level 3. Then, each children subtree is coded as what has been carried on its parent. In the end, every children subtree will find a domain subtree which has the minimum approximation distortion. This variable tree size fractal coding algorithm can be repeatedly applied on the children subtrees until every range subtree has found its best matched domain subtree within given distortion, or the allowed minimum dimension of range subtree has been reached which means it can not be further split and will find the best-matched domain subtree regardless of the distortion.

III. WAVELET SUBTREE BASED FRACTAL VIDEO CODING

Assuming that a video frame is transformed to wavelet domain by a 4 stage pyramidal decomposition, the motion activities at different layers of the wavelet pyramid are different but highly correlated since they actually characterize the same motion structure. Thus, the wavelet coefficients at different scales and different frequency bands also can be grouped together to form the wavelet subtree which gives a new way to represent the motion structure. As the proposed video coding scheme is to exploit the redundancy among adjacent frames not the redundancy within the frame itself, the range subtrees are formed by the coefficients from the current encoded frame, but the domain subtrees are constructed by the coefficients from the previous frame instead of the current frame. Moreover, the lowest frequency subband is considered in the range and domain subtree construction such that the highest layer of subtree contains only single root node which has three children nodes from the three high frequency subband at the same scale. The range and domain subtrees now have the same total number of tree nodes which is more convenient for matching comparison.

For better coding performance, not all range subtrees are coded with same effort, it is because that some range subtrees may have no motion information which can be predicted by the corresponding domain subtree with same spatial location in previous frame. Thus a range subtree is first judged by a proper multiresolution motion detection algorithm as described in the next section.

1. Multiresolution motion detection (MRMD)

The proposed MRMD is based on the mean squared error (MSE) between a range subtree and its corresponding domain subtree with same spatial location in previous frame. If the MSE between them is below a predefined threshold, the range subtree is a non-motion subtree. Otherwise, it is a motion subtree. Then coding requirement for a non-motion range subtree is reduced to the minimum that only a terminal flag is assigned to the whole subtree which is sufficient for reconstruction when the previous frame is given. On the contrary, if a range subtree contains more motion information and can not be simply approximated by its predecessor, an efficient coding algorithm must be followed.

2. Variable tree size fractal coding of motion subtrees

The motion subtree is first segmented into two parts. The one part only contains the root node and will be scalar quantized within its own subband; the other part consists of the rest tree nodes which is taken as a new wavelet subtree.

All the initial domain subtrees in previous frame also minus the root node to form the corresponding searching pool. Now the main task of the proposed fractal motion coder is to find the best-matched domain subtree to approximate the given motion subtree through proper affine transformation. Then, the matching distortion ε is compared to a target distortion T_s. If ε is not larger than T_s, the range subtree is coded using "fractal quantization" and further splitting is not necessary. On the contrary, if ε exceeds T_s, the range subtree will be segmented into several parts because the current error is too large to allow a good approximation.

Similarly, the three root nodes are removed from the tree and each of them is coded using pixel-based scalar quantization within its own subband. The rest nodes are split into four quadrant children subtrees with the root at next finer scale level 3. The domain subtrees are constructed such that their root nodes at the same scale level. Each children subtree is encoded as its parent. This adaptive fractal coding for motion subtrees will be repeated until every children subtree has found its best matched domain subtree within predefined distortion threshold, or it has reached the minimum size of range subtree and hence further splitting is not allowed. In this case, the best-matched domain subtree with minimum distortion is found even such approximation is poor.

3. Overall structure

The proposed new wavelet transform based fractal video compression scheme can be divided into two parts: intraframe and interframe coding as illustrated in Fig. 2. The input sequence is first transformed into wavelet domain and followed by choice of intra or inter frame coding. The first frame of video sequence is always encoded in intraframe mode by using the variable wavelet subtree size based fractal coding as described in section II. To code the successive frames, the first frame is reconstructed in the wavelet domain by fractal decoding algorithm and then put in the frame memory as reference.

In our simulation, the remaining frames in the video sequence are all coded in interframe mode. It begins from MRMD, in which all range subtrees of the current frame are classified into two categories, motion and non-motion subtrees. The encoding of non-motion subtree is very simple, while the motion subtrees are coded by variable tree size fractal-based coding algorithm as described above. Because the domain subtrees are extracted from previous frame (frame memory) not from current frame itself, only one step fractal decoding is enough to obtain the approximation of original frame. The frame memory is refreshed periodically with the newly decoded frame before coding the next frame. The encoding, decoding and refreshing process is repeated for each of the left frames. The original video frame can be obtained from applying inverse wavelet transform on the coefficients which are put in the frame memory. The whole video sequence will be recovered sequentially.

IV. EXPERIMENTAL RESULTS

Computer simulations are carried out on standard QCIF image sequences with 176×144 pixels in each frame. Only the luminance part of pixels is considered in the implementation. However, it can be easily extended for color video sequence. The two test sequences are "Miss America" and "Salesman" with 150 frames each. The biorthogonal wavelets with less dissimilar lengths of filters "9-7", referred as B97, are employed in wavelet transform [1]. Each video frame is transformed to wavelet domain by a 4 scale pyramidal decomposition producing a total of 13 subbands. In intraframe mode, the minimum size of range subtree is 15 which means only one stage splitting is used. In interframe coding mode, the search region is restricted to -3~+3 in both horizontal and vertical direction to reduce the encoding complexity. The minimum size of range subtree is also 15, thus 3 stage splitting for a motion subtree is employed. At the end of the video encoder, adaptive arithmetic coding is applied to generate the output bit stream as the coding results. The most commonly used measure PSNR is given as an indication of the image quality of reconstructed frame. It is defined as PSNR = $10\log_{10}(255^2/\text{MSE})$, where the MSE is the mean squared error between the original and decoded frame.

The proposed video codec has obtained the bit-rates of 0.056bpp and 0.110bpp with the averaged PSNR 33.51dB and 31.14dB for the test sequence "Miss America" and "Salesman", respectively. In general, the "Salesman" has more motion activities than the "Miss America" and the background of the former is more complicated than that of the later. These are the reasons that "Miss America" has more averaged non-motion subtrees and hence obtain a higher compression ratio than the "Salesman". The PSNR for each reconstructed frame is shown in Fig. 3. We can see that the PSNR of the two test sequences are relatively constant due to less propagation error incurred in the coding scheme. Two reconstructed frames from different video sequence are shown in Fig. 4. As it may be seen, no blocking effects can be observed which is a big problem of block-based fractal video coding scheme. However, we can detect some artifacts and smearing effects due to the quantization error.

V. CONCLUSIONS

A new wavelet transform based fractal video compression scheme is proposed in this paper. It exploits the correlation between wavelet subtrees not only within the video frames but also among the adjacent frames. One of the most advantages using wavelet subtree representation of motion structure is that it can describe large object motion activities instead of piecewise motion activities [2]. Therefore, it requires fewer computation since separate parts of wavelet subtree are processed together if their own motion activities coincide with the global motion information. This is also the reason that it can obtain lower bit-rates because separate pieces of a wavelet subtree are now coded as a whole. Moreover, because the blocking artifacts are reduced significantly, the visual quality of the reconstructed frames is better as compared to those using fractal video coding in spatial domain.

REFERENCES

[1] M. Antonini, M. Barlaud, P. Mathieu and I. Daubechies, "Image coding using wavelet transform," *IEEE Trans. on Image Proc.*, vol. 1, no. 2, pp. 205-220, Apr. 1992.

[2] Y. Q. Zhang and Z. Zafar,, "Motion-compensated wavelet transform coding color video compression," *IEEE Trans. Circuit & Systems for Video Technology*, vol. 2, no. 3, pp. 285-298, Sep, 1992.

[3] A. Jacquin, "Fractal image coding based on a theory of iterated contractive image transformations," in *Proc. SPIE Visual Comm. and Image Proc.*, vol. 1360, pp. 227-239, 1990

[4] Y. Fisher, "Fractal image compression with quadtrees," in *Fractal Compression: The ory and Application to Digital Images*, Y. Fisher, Ed. Springer Verlag, New York, 1994.

[5] M. S. Lazar and L. T. Bruton, "Fractal block coding of digital video," *IEEE Trans. on Circuits & Systems for Video Technology*, vol. 4, no. 3, pp. 297-308.

[6] G. Davis, "Self-quantization of wavelet subtrees," in *Proc. SPIE Wavelet Applications II*, Orlando, vol. 2491, pp.141-152, Apr. 1995.

[7] H. Krupnik, D. Malah, and E. Karnin, "Fractal representation of images via the discrete wavelet transform," in *IEEE 18th Conv. of EE in Israel*, Tel-Aviv, Mar. 1995.

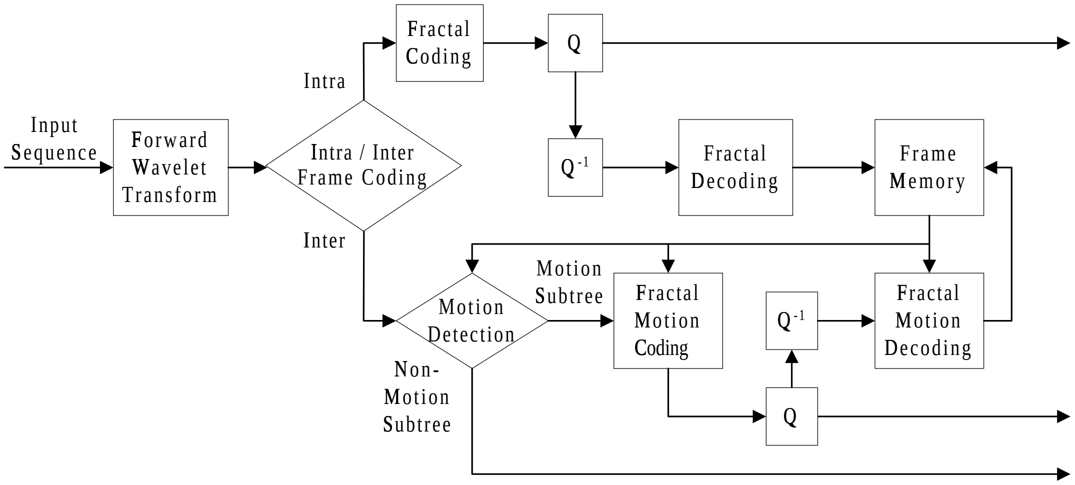

Fig. 2 Block diagram of the wavelet transform based fractal video coding scheme

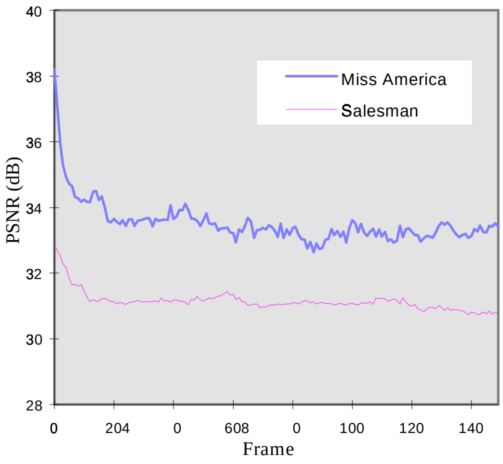

Fig. 3 PSNR versus frames for test sequences.

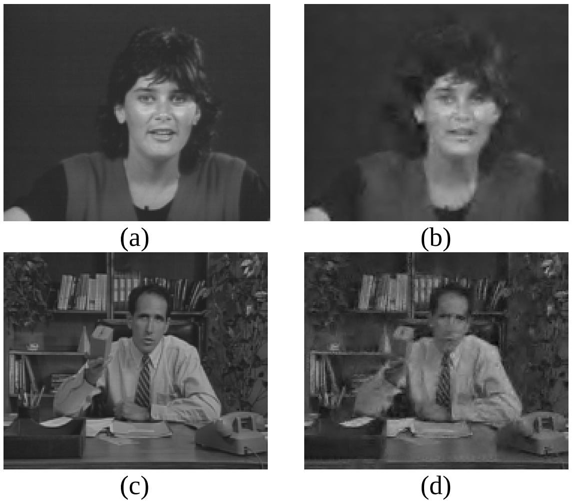
(a) (b)
(c) (d)

Fig. 4. Decoded frame 81 of the test sequence
(a) "Miss America", original (b) Decoded frame, PSNR = 33.38dB
(c) "Salesman", original (d) Decoded frame, PSNR = 31.10dB.