FRACTAL CODING OF VIDEO SEQUENCE BY CIRCULAR PREDICTION MAPPING

CHANG-SU KIM and SANG-UK LEE
Inst. of New Media and Communications
Seoul National University
Seoul 151-742, Korea
E-mail : cskim@Claudia.snu.ac.kr

Abstract

We propose a novel algorithm for fractal video sequence coding, based on the Circular Prediction Mapping (CPM), in which each range block is approximated by a domain block in the circularly previous frame. In our approach, the size of the domain block is set to be same as that of the range block for exploiting the high temporal correlation between the adjacent frames, while most other fractal coders use the domain block larger than the range block. Therefore the domain-range mapping in the CPM becomes similar to the block matching algorithm in the motion compensation techniques, and the advantages from this similarity are discussed. Also the modified collage theorem, which yields better prediction method for the CPM than the conventional collage theorem, is provided by linear systematic analysis. The computer simulation results on real video-conferencing image sequences demonstrate that the average compression ratios ranging from 60 to 70 can be obtained with good subjective quality.

1. INTRODUCTION

Fractal compression, which is based on the IFS (iterated function system) proposed by Barnsley[1], is a new approach to image coding recently. The basic notion of the fractal image compression is to find a contraction mapping whose unique attractor approximates the source image. In the decoder, the mapping is applied iteratively to an arbitrary image

to reconstruct the attractor. If the mapping can be represented with less bits than the source image, a coding gain is obtained.

After Jacquin[2] proposed the first automatic algorithm for fractal coding of still images, much effort[3,4,5,6] has been made to the fractal still image coding techniques. However, little work has been reported on the fractal video sequence coding techniques. Lazar[7] and Li[8], respectively, extended the still image coding techniques straightforwardly to the video sequence coding, by employing the 3-D domain blocks and range blocks. The Lazar's algorithm is the extended version of the Jacquin's algorithm[2] and the Li's algorithm is the extended version of the Monro's algorithm[6], respectively. But these algorithms are very complicated to implement, and severe 3-D blocking artifacts are observed in the reconstructed images in many cases. Alternative approach, which encodes each frame using the previous frame as a domain pool, was proposed by Fisher[9]. The main advantage of this algorithm is that fast decoding is possible, since it does not require iteration at the decoder. But, in this algorithm, the temporal correlation between the frames may not be effectively exploited, since the size of the domain block is larger than that of the range block.

In this paper, a novel approach, called the Circular Prediction Mapping (CPM), is proposed to combine the fractal sequence coder with the well-known motion estimation/motion compensation (ME/MC) techniques, so that the proposed algorithm is capable of exploiting the high temporal correlation between the frames. In the CPM, n frames are encoded as a group, and each range block is motion-compensated by a domain block in the n-circularly previous frame, which is of the same size as the range block. By selecting appropriate parameters in the domain-range mappings, the CPM becomes a contraction mapping. In the decoder, the CPM is applied iteratively to arbitrary n frames to reconstruct the attractor frames.

The collage theorem[1] provides the attractor error bound at the decoder, in terms of the collage error at the encoder. But we shall show that the attractor error bound can be significantly reduced by modifying the collage theorem, due to the fact that the search region for the domain block is confined to the n-circularly previous frame in the CPM. This modified collage theorem gives a new prediction method at the encoder. The computer simulation results on real video-conferencing image sequences demonstrate that the average compression ratios ranging from 60 to 70 can be obtained with the good subjective quality, by employing this new prediction method for the CPM.

2. BACKGROUND

The fractal image compression techniques are based on *the contraction mapping theorem* and *the collage theorem* [1]. In this section, we briefly describe these two theorems for the sake of completeness.

Let $(X, \|\cdot\|)$ be the complete metric space of still images or video sequences. Then a transformation $f : X \rightarrow X$ is called a contraction mapping, if there exists a constant $0 \leq s < 1$ such that

$$\|f(x) - f(y)\| \leq s \cdot \|x - y\| \qquad \text{for all } x, y \in X. \tag{1}$$

The constant s is called the contractivity factor for f. *The contraction mapping theorem* ensures that each contraction mapping f has a unique attractor (fixed point) x_f, such that

$$f(x_f) \;=\; x_f. \tag{2}$$

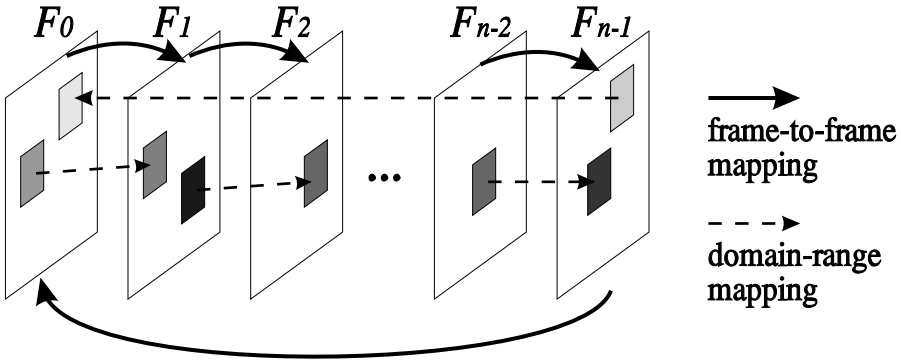

Figure 1: The structure of the circular prediction mapping (CPM)

Moreover, the f can be applied iteratively to an arbitrary point $y \in X$ to obtain the attractor x_f, by

$$\lim_{n\to\infty} f^n(y) \quad = \quad x_f. \tag{3}$$

In the context of image coding, if the encoder find a contraction mapping whose unique attractor is the source image, then the mapping can be applied iteratively to an arbitrary image to reconstruct the source image in the decoder. Therefore a coding gain is obtained, if the mapping can be represented with less bits than the source image. But it is practically impossible to find a contraction mapping whose attractor is exactly the source image, since it requires too many calculations and the size of the encoded data for the mapping gets very large as we attempt to encode the source image losslessly.

As a lossy coding technique, the fractal encoder attempts to find the contraction mapping f whose collage $f(x)$ is close to the source image x. Then *the collage theorem* provides the relation between the collage error at the encoder $\|x - f(x)\|$ and the attractor error at the decoder $\|x - x_f\|$, given by

$$\|x - x_f\| \leq \frac{1}{1-s} \cdot \|x - f(x)\|, \tag{4}$$

where s is the contractivity factor for f. This means that the decoded attractor x_f is close to the source image x, if the collage $f(x)$ is close to the source image x. The optimization of the collage, instead of the attractor, alleviates high complexity of the encoder.

3. CIRCULAR PREDICTION MAPPING (CPM)

In this section, the proposed CPM, which is a suitable contraction mapping for encoding and decoding of moving image sequences, is described and many properties relating to the CPM are discussed.

In the CPM, n frames are encoded as a group, and each frame is predicted blockwise from the n-circularly previous frame, as shown in Figure 1. The k-th frame F_k is partitioned into the range blocks, and each range block in F_k is predicted or approximated by a domain block in $F_{[k-1]_n}$, where $[k]_n$ denotes (k *modulo* n).

The size of the domain block is set to be same as that of the range block in our approach, while the domain block is larger than the range block in most other fractal coders[2,7,8]. Since

there exist many temporal redundancies in moving image sequences, there is more chance of good domain-range mapping when the size of the domain block is same as that of the range block, in which case the algorithm to find the optimal domain block becomes similar to the block matching algorithm in the motion compensation techniques.

3.1 Mathematical Model of the CPM

To the end of this section, we assume without loss of generality that 4 frames are encoded as a group, *i.e.*, the length of a coding group n is 4. Then the CPM is composed of 4 frame-to-frame mappings, and each frame-to-frame mapping is sum of domain-range mappings.

• **domain-range mapping** : Each range block R_i in the k-th frame F_k is approximated by a domain block $D_{a(i)}$ in the 4-circularly previous frame $F_{[k-1]_4}$, which is of the same size as the range block. The approximation of the R_i is given by

$$R_i \cong \widetilde{R_i} = s_i \cdot \mathcal{O}(D_{a(i)}) + o_i \cdot C, \tag{5}$$

where $a(i)$ denotes the location of the optimal domain block, and s_i, o_i are real coefficients, respectively. C is a constant block whose all pixel values are 1, and $\mathcal{O}$ is the orthogonalization operator which is proposed by Øien[12] for fractal still image coder. This operator removes DC component from $D_{a(i)}$, so that $\mathcal{O}(D_{a(i)})$ and C are orthogonal to each other. After the orthogonalization, the optimal coefficient values of s_i, o_i can be directly obtained by projection of R_i onto $span\{\mathcal{O}(D_{a(i)})\}$ and $span\{C\}$, respectively. Notice that the s_i coefficient determines the contrast scaling in the mapping, and the o_i coefficient represents the DC value of the range block R_i.

This domain-range mapping can be interpreted as a kind of motion compensation techniques. We adopt a motion model that describes the motion of image objects by the translationally moving blocks. Then the $a(i)$ describes the translational motion of block, *i.e.*, the $a(i)$ is the motion vector. Besides the translational motions, the changes in contrast and overall brightness of blocks are compensated by the s_i and o_i coefficients, respectively.

The main advantage of the proposed domain-range mapping can be realized from the above interpretation. In real moving image sequences, small motion vector is more probable than large motion vector. Therefore the search region for the motion vector $a(i)$ can be localized to the area near the location of the range block, while most other fractal coders should search over much larger region to find a good domain-range mapping, alleviating the computational burden of the encoder. In addition, the $a(i)$ can be coded with less bits.

• **frame-to-frame mapping** : The frame-to-frame prediction mapping from the $F_{[k-1]_4}$ to the F_k is sum of the domain-range mappings in Eq.(5), given by

$$F_k \cong F_k^c = L_k F_{[k-1]_4} + T_k, \qquad (0 \le k < 4) \tag{6}$$

where F_k^c denotes the k-th collage (predicted) frame, L_k is the appropriately rearranged square matrix which is composed of the s_i coefficients and the orthogonalization operators in Eq.(5), and T_k is the rearranged column matrix which is composed of the o_i coefficients in Eq.(5), respectively.

• **circular prediction mapping** : The CPM is composed of the 4 frame-to-frame mappings in Eq.(6), given in array form by

$$\boldsymbol{F}^c \;=\; \boldsymbol{L}\boldsymbol{F} + \boldsymbol{T}$$

$$= \begin{bmatrix} O & O & O & L_0 \\ L_1 & O & O & O \\ O & L_2 & O & O \\ O & O & L_3 & O \end{bmatrix} \boldsymbol{F} + \begin{bmatrix} T_0 \\ T_1 \\ T_2 \\ T_3 \end{bmatrix}, \quad (7)$$

where $\boldsymbol{F} = [F_0, F_1, F_2, F_3]^T$ denotes the source sequence, and $\boldsymbol{F}^c = [F_0^c, F_1^c, F_2^c, F_3^c]^T$ denotes the collage sequence, respectively. This CPM is an affine mapping which is determined by the linear operator $\boldsymbol{L}$ and the translation vector $\boldsymbol{T}$.

In the domain-range mapping, the DC value for each range block is represented by the o_i coefficient, and the other AC informations are contained in the motion vector $a(i)$ and the contrast scaling factor s_i. Therefore by inspection of Eqs.(5)-(7), it can be observed that the DC and AC informations for all the range blocks are contained in the translation vector $\boldsymbol{T}$ and the linear operator $\boldsymbol{L}$, respectively.

In the decoder, the CPM is applied iteratively to an arbitrary sequence to obtain the attractor sequence. Thus the CPM should be contractive for the iterative process to converge. Since the CPM is an affine mapping, the contractivity factor for the CPM is given by the norm of the linear operator $\|\boldsymbol{L}\|$, which is the square root of the largest eigenvalue of $\boldsymbol{L}^T\boldsymbol{L}$ in l^2-metric [10]. Therefore the encoder should select appropriately the coding parameters $a(i)$, s_i in Eq.(5), which determine the linear operator $\boldsymbol{L}$, so that all the eigenvalues of $\boldsymbol{L}^T\boldsymbol{L}$ are less than 1. But it is practically impossible to compute the eigenvalues of $\boldsymbol{L}^T\boldsymbol{L}$ everytime we select or quantize the coding parameters, since the matrix $\boldsymbol{L}$ is very large and complicated in general. To overcome this problem, we constrain the contrast scaling coefficients s_i to be quantized between -1 and 1 at the encoder, instead of controlling the eigenvalues of $\boldsymbol{L}^T\boldsymbol{L}$ directly. This constraint ensures the contractivity of the domain-range mapping given in Eq.(5), since the contractivity factor of the domain-range mapping is given by the norm of the linear operator $\|s_i\mathcal{O}\|$ and the orthogonalization operator $\mathcal{O}$ always has unit matrix norm. In other words, the local contractivity of the CPM is guaranteed by the constraint on the contrast scaling coefficients s_i. Then the iterative applications of the CPM will be *eventually contractive*[3] in most cases, even if the contractivity factor of the CPM is larger than 1.

3.2 Modified Collage Theorem for the CPM

The collage theorem gives the attractor error bound at the decoder, in terms of the collage error at the encoder, as mentioned in section 2. Most fractal encoders attempt to find the contraction mapping, which minimizes the collage error, based on this theorem. Figure 2.(a) illustrates the direct application of the collage theorem to the proposed CPM, in which the encoder minimizes the prediction error (collage error) of F_k from the 4-circularly previous frame $F_{[k-1]_4}$. In other words, each frame is predicted from the 4-circularly previous source frame at the encoder.

But the attractor error bound can be reduced by modifying the collage theorem without increasing the complexity of the encoder, since the search region for the domain block is confined to the 4-circularly previous frame in the proposed CPM. The modified collage theorem yields a new prediction method, and the new prediction method is illustrated in Figure 2.(b). Firstly, F_0 is predicted from F_3. Secondly, F_1 is predicted from the previously predicted frame F_0^c, rather than the previous source frame F_0. Similarly, F_2 and F_3 are predicted from the previously predicted frames $(L_1F_0^c + T_1)$ and $[L_2(L_1F_0^c + T_1) + T_2]$, rather

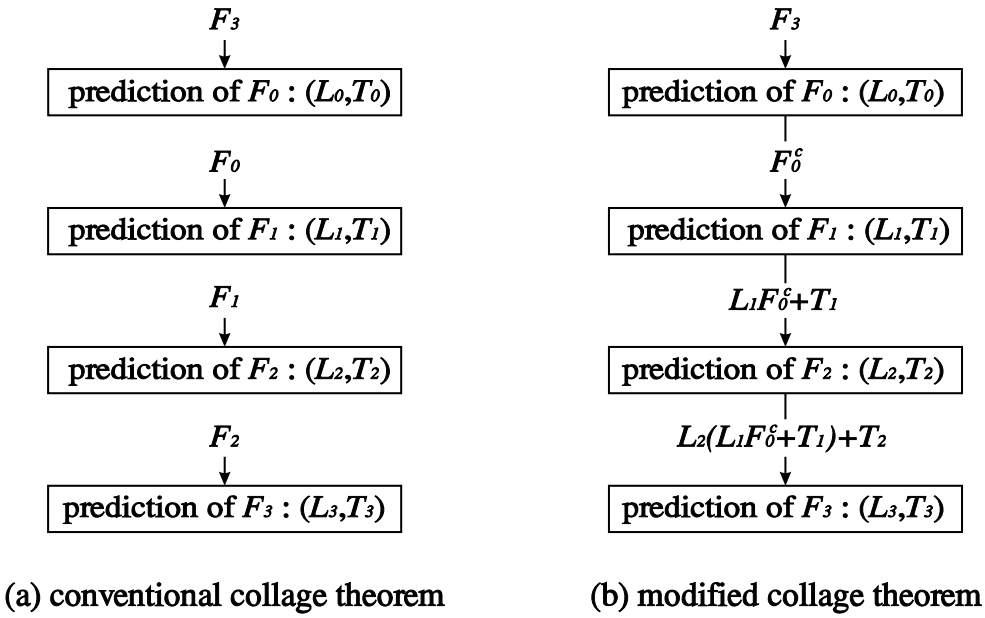

Figure 2: The comparison of the Prediction Methods

than the previously source frames F_1 and F_2, respectively. Then the attractor error bound of the modified collage theorem can be shown to be less than that of the conventional collage theorem. The rigorous proof of the modified collage theorem is provided in Appendix.

Let us discuss the intuitive interpretation of the modified collage theorem. If we assume that F_3 is reconstructed without any distortion at the decoder, then it is clear from Figure 2 that the modified collage theorem will reconstruct the other frames F_k $(0 \leq k < 3)$ more closely to the original frames than the conventional collage theorem, since the modified collage theorem prohibits the error propagation by employing the previously predicted frame, rather than the previous source frame, as a domain pool. This argument is very analogous to the closed-loop DPCM (Differential Pulse-Code Modulation) [11], in which the prediction depends on the previously quantized values, rather than on the previously unquantized values, to prevent the propagation of the quantization error. Even if there is some distortion in the reconstructed F_3, this prevention of the error propagation yields smaller attractor error bound than the conventional collage theorem, which does not consider the error propagation along the frames.

3.3 Decoding Algorithm

The attractor sequence can be reconstructed by iteratively applying the CPM to an initial arbitrary sequence. In general, the convergence speed is dependent on the ratio of the size of the domain block and the size of the range block. The larger is the domain block as compared to the range block, the faster the decoded sequences converge. But in our CPM, the size of the domain block is set to be same as that of the range block to exploit the temporal correlation, and the convergence speed is very slow.

But this disadvantage is compensated by the advantage of the CPM that one iteration of the CPM has the effect of 4 (= the length of a coding group) iterations in other fractal coders. This merit is due to the fact that the search region for the domain block is confined to

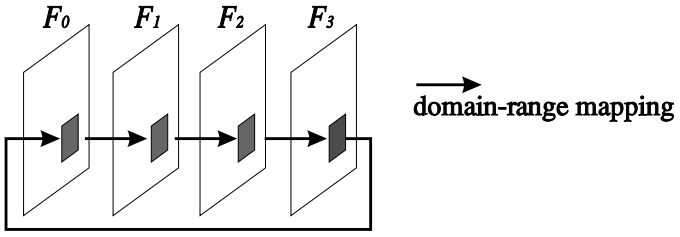

Figure 3: The Illustration of Self-Mapping Effect

the 4-circularly previous frame. Let F_k^i denote the k-th decoded frame at the i-th iteration. At the first iteration, F_0^1 is decoded by applying (L_0, T_0) to an arbitrary frame, and F_k^1 $(1 \leq k < 4)$ is subsequentially decoded by applying (L_k, T_k) to the previously decoded frame F_{k-1}^1, respectively. Note that F_3^1 is more closer to the attractor sequence than F_0^1, since F_3^1 is actually the result of 4 iterations. At the second iteration, F_0^2 is decoded using F_3^1 of the first iteration, and so on. This process is repeated until the difference between the outputs from successive iterations becomes small.

4. DESIGN OF CPM FRACTAL CODER

There exist many variations in design or implementation of CPM fractal sequence coder. For example, image partitioning scheme should be established before the implementation. This section describes several issues relating to the implementation of the CPM fractal coder in more detail.

4.1 Image Partitioning

Since moving image sequences are unbounded in temporal direction, they should be temporally partitioned before encoding. In our implementation, each 4 frames of input sequences are encoded as a coding group ($n = 4$). If the length of a coding group n is large, the backward prediction error of F_0 from F_{n-1} will become larger and there will be too much time-delay between the encoder and the decoder. On the contrary, if the n is small, the overall bit-rate will become higher (this will be explained in 4.3) and the decoding speed will become slower. Therefore, the n is selected to be 4 as a tradeoff.

After the temporal partitioning, each frame is spatially partitioned into the range blocks of maximum size 32×32 and minimum size 4×4, using the quadtree structure [13]. First, 32×32 range block is approximated by a domain block in the 4-circularly previous frame, and if the approximation error is larger than the pre-specified threshold, then it is decomposed further into four smaller 16×16 range blocks. This process is repeated until the approximation error is smaller than the threshold or 4×4 range block is generated.

4.2 Elimination of Self-Mapping Effect

An undesirable effect, called self-mapping effect, may be experienced in the CPM structure, since the size of the domain block is same as that of the range block. Figure 3 illustrates the case when the self-mapping effect happens. Each of the four range blocks in Figure 3 is approximated by itself. In other words, the four range blocks are self-mapped when

the decoder applies the CPM four times. The AC informations of these range blocks are approximated by the contrast-scaled AC informations of themselves, and converge to zero at the decoder since the contrast scaling factors s_i's in Eq.(5) are quantized between -1 and 1. Therefore these range blocks are reconstructed as constant blocks, which have only the DC informations, yielding severe blocking artifacts in the decoded frames.

This self-mapping effect is eliminated compulsively in our implementation by disjointing the address of the optimal domain block in the forward domain-range mappings from that in the backward domain-range mappings. More specifically, the address $a(i)$ for the range block in F_k $(1 \leq k < 4)$ is constrained to be even numbers in both the x,y coordinates, but the address $a(i)$ for the range block in F_0 is constrained to be odd numbers in both the x,y coordinates.

4.3 Parameter Quantization and Bit Allocation

For efficient transmission or storage, it is necessary to quantize the coefficients[4]. Let us describe the issue relating to the quantization in more detail.

The compressed data for each range block are composed of the address $a(i)$ and the s_i, o_i coefficients. Firstly, the s_i coefficients representing the contrast-scaling in the domain-range mappings are fixed as 0.9, since they are distributed compactly at the center of 1 and should be quantized between -1 and 1 to ensure the contractivity of the CPM. Therefore no bit is allocated to the s_i coefficients. Secondly, the o_i coefficient is the DC value of the range block R_i, and is highly correlated with the DC value of the optimal domain block $D_{a(i)}$. Thus the o_i in F_k $(1 \leq k < 4)$ is predicted from the o_i's in F_{k-1}, and the prediction error is coded with the Huffman coder. But the o_i in F_0 is uniformly quantized with 8 bits between $0 \sim 255$ for the causality of the system, so the bit-rate for F_0 is usually higher than the bit-rates for the other frames in the same coding group. Lastly, the search region for the optimal domain block $D_{a(i)}$ is the rectangular area centered at the location of the range block, and the address $a(i)$ is expressed with respect to the location of the range block. The coordinates of the $a(i)$ in the forward and backward domain-range mappings are constrained to be in $\{-16, -14, \ldots, 12, 14\}$ and $\{-15, -13, \ldots, 13, 15\}$, respectively. Therefore 8 bits are needed for representing each $a(i)$ with fixed-length codewords. But the probability distribution of the $a(i)$ is not uniform, since the $a(i)$ is the counterpart of the motion vector in ME/MC. More specifically, the small motion vector is more probable than the large motion vector. This non-uniformity is exploited by the Huffman coder, using the probability distribution function obtained from many test sequences.

5. SIMULATION RESULTS

The proposed algorithm is tested on real moving image sequences. As illustrated in Figure 2, two different prediction methods for the CPM fractal encoder could be considered. The first method is the direct application of the conventional collage theorem, and the second prediction method is based on the modified collage theorem. Figure 4 shows the performances of these two prediction methods on the standard CIF (352×288) "Miss America" sequence. It can be seen that the second method provides better PSNR (about $1 \sim 2$ dB) performance than the first method. Notice that the average bit-rate for the first prediction method (= 0.1242 bpp) is slightly higher than that of the second method (=0.1235

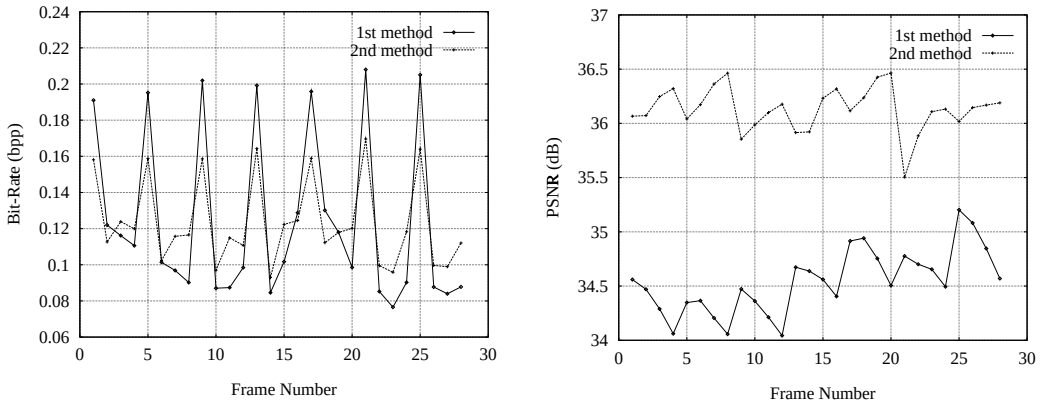

Figure 4: The bit-rate and PSNR curves for "Miss America" sequence

bpp). These simulation results coincide with the analytical approach in Section 3, which shows that the attractor error bound of the modified collage theorem is less than that of the conventional collage theorem. Thus only the second prediction method is employed in the following simulations.

In Figure 5, we present an example of the decoded results of the first iteration for "Miss America" sequence. This example is for a coding group which are composed of the 1st $\sim$ 4th frames. Let F_k^i denote the k-th decoded frame at the i-th iteration, as in 3.3. At first, F_0^1 is decoded by applying the (L_0, T_0) to an initial uniformly gray image, as shown in the upper left corner of Figure 5. It can be observed that each range block is reconstructed as a constant block whose all pixel values are the DC value of the range block, since the orthogonalization operator is incorporated into the domain-range mapping and the domain block from the uniformly gray image has no AC information. Also notice that the smooth region such as the background is partitioned into large range blocks, and the detailed region such as the face is partitioned into small range blocks, respectively, according to the quadtree structure. F_1^1 is decoded by applying the (L_1, T_1) to F_0^1, and F_1^1 is actually the result of two iterations. Likewise, F_2^1 and F_3^1 are actually the result of three and four iterations, respectively. Therefore it is observed that F_3^1 is much more detailed than F_0^1. At the second iteration, F_0^2 is decoded by employing F_3^1 as a domain pool, and so on. It has been observed that at most five iterations of the CPM, which is actually 20 iterations, is sufficient for the decoded frames to converge in most cases.

Figure 6 shows the bit-rate and PSNR performances for other various CIF (352×288) image sequences. From Figure 4 and 6, it can be seen that the bit-rate for F_0 is higher than the bit-rates for the other frames (F_1, F_2, F_3) in the same coding group. This is due to the fact that the o_i coefficients of F_0 are encoded with FLC (variable length codewords), while the o_i coefficients of the other frames are predicted from the previous frame and the prediction errors are encoded with VLC (variable length codewords). Moreover, F_0 is backwardly predicted from F_3 which are 3 frames apart, while the other frames are forwardly predicted from the previous frame. Therefore F_0 is less efficiently predicted than the other frames, and is partitioned into smaller range blocks. But the quality of reconstructed frames is observed to be almost comparable in a coding group, since each frame of a coding group is employed as a domain pool for the circularly next frame and the error of a frame spreads into the other frames.

For the "Miss America" and "Claire" sequences, the average bit-rates are 0.124 and 0.116

upper left : F_0^1 (1st frame) upper right : F_1^1 (2nd frame)
lower left : F_2^1 (3rd frame) lower right : F_3^1 (4th frame)

Figure 5: The first iteration for "Miss America" 1-4

bpp, yielding 64.5 and 68.9 of the compression ratios, respectively. In other words, $5 \sim 6$ frames can be transmitted in a second at the bandwidth of 64 kbits/s. The bit-rates for the "Foreman" and "Car Phone" sequence are higher than those of the "Miss America" and "Claire" sequences. This is inevitable since these sequences are fine-detailed and contain large inter-frame motions. Fig.7 presents samples of the decoded frames. It is observed that the proposed algorithm reconstruct the "Miss America" and "Claire" sequences with a good subjective quality which is sufficient for the video-conferencing applications, since it does not yield severe blocking artifacts, which are main defects of the 3-D block approaches[7,8]. It is also observed that the tree outside the window in the "Car Phone" sequence and the sharp edge of the background structure in "Foreman" sequence are reconstructed very faithfully, though they move very fast. This is due to the fact that the domain-range mapping of the CPM is very similar to the ME/MC techniques. In the 3-D block approaches, the domain-range mapping often fails, and the quality of the reconstructed frames is poor in such finely-detailed and fast moving regions. These simulation results indicate that the proposed algorithm provides much better performance than the conventional 3-D block approaches.

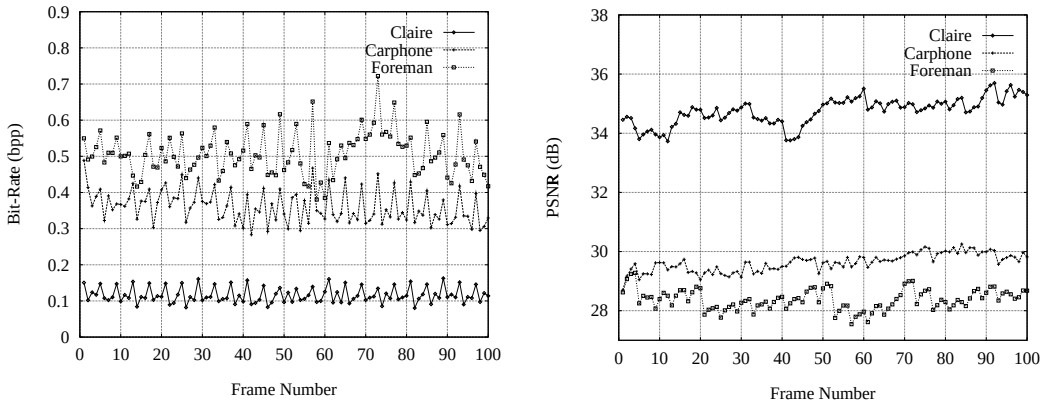

Figure 6: The bit-rate and PSNR curves for the various CIF sequence

6. CONCLUSION

In this paper, we proposed a novel algorithm for fractal video sequence coding, based on the Circular Prediction Mapping (CPM). In our approach, each range block was approximated by a domain block in the 4-circularly previous frame, and the size of the domain block was set to be same as that of the range block, in order to exploit the high temporal correlation in real moving image sequences. By linear systematic analysis, we derived the modified collage theorem, which yields better prediction method than the conventional collage theorem.

It was demonstrated by the computer simulation on the "Miss America" and "Claire" sequences that the average compression ratios ranging from 60 to 70 can be achieved without observing severe blocking artifacts in the reconstructed sequences, while the other fractal sequence coders generate severe 3-D blocking artifacts at such low bit-rates.

Further research will include the following issues to improve the performance of the CPM fractal sequence coder further.

- Adaptive grouping of frames – In this paper, the length of a coding group n is fixed as 4. But the n should be varied, according to the characteristics of the frames. For example, the coding group should not contain a scene change in it.

- Inclusion of *condensation sets*[1] – Intra-coded block is called a condensation set in fractal terminology. By including the condensation sets where the domain-range mapping fails, the performance of the fractal coder will be improved. The condensation sets may be coded with other coding techniques, such as DCT.

- Optimization of the coding parameters – The coding parameters, proposed in this paper, are not optimal at all. More experiments are needed to find the relation between the performance of the CPM fractal coder and the selected parameter sets.

We believe that this research will increase the coding performance significantly, and the CPM fractal coder will be a strong candidate for the very low bit-rate coding techniques.

7. REFERENCES

1. M. F. Barnsley, *Fractals Everywhere* (Academic Press, San Diego, 1988).

upper left : Miss America upper right : Claire
lower left : Car Phone lower right : Foreman

Figure 7: The samples (1st frames) of the decoded sequences

2. A. E. Jacquin, *Image Coding Based on a Fractal Theory of Iterated Contractive Image Transformations.* IEEE Trans. on Image Processing, **Vol.1, No.1**, p.18-30, January 1992.
3. Y. Fisher (editor), *Fractal Image Compression-Theory and Application* (Springer-Verlag, New York, 1995).
4. G. E. Øien, *Parameter Quantization in Fractal Image Coding.* ICIP, **Vol.3**, p.142-146, 1994.
5. S. Lepsøy, G. E. Øien and T. A. Ramstad, *Attractor Image Compression with a Fast Non-iterative Algorithm.* ICASSP, **Vol.5**, p.337-340, 1993.
6. D. M. Monro and F. Dudbridge, *Fractal Approximation of Image Blocks.* ICASSP, **Vol.3**, p. 485-488, 1992.
7. M. S. Lazar and L. T. Bruton, *Fractal Block Coding of Digital Video.* IEEE Trans. on Circuits and Systems for Video Technology, **Vol.4, No.3**, p.297-308, June 1994.
8. H. Li, M. Novak and R. Forchheimer, *Fractal-based Image Sequence Compression Scheme.* Optical Engineering, **Vol.32, No.7**, p.1588-1595, 1993.
9. Y. Fisher, T. P. Shen and D. Rogovin, *Fractal (Self-VQ) Encoding of Video Sequences.* SPIE:VCIP, **Vol.2304-16**, 1994.
10. A. W. Naylor and G. R. Sell, *Linear Operator Theory in Engineering and Science* (Springer-Verlag, New York, 1982).
11. N. S. Jayant and P. Noll, *Digital Coding of Waveforms - Principles and Applications to Speech and Video* (Prentice-Hall, New Jersey, 1984).

12. G. E. Øien, S. Lepsøy and T. A. Ramstad, *An Inner Product Space Approach to Image Coding by Contractive Transformations.* ICASSP, p.2773 - 2776, 1991.
13. E. Shusterman and M. Feder, *Image Compression via Improved Quadtree Decomposition Algorithms.* IEEE Trans. on Image Processing, **Vol.3, No.2**, p. 207-215, March 1994.

APPENDIX - Proof of Modified Collage Theorem

According to the Contraction Mapping Theorem, there exist unique attractor sequence $\boldsymbol{F}^o = [f_0^o, f_1^o, f_2^o, f_3^o]^T$, which is determined by the CPM $(\boldsymbol{L}, \boldsymbol{T})$, such that

$$\boldsymbol{F}^o = \boldsymbol{L}\boldsymbol{F}^o + \boldsymbol{T}. \tag{8}$$

The attractor error $\boldsymbol{E}$ at the decoder and the collage error $\boldsymbol{E}^c$ at the encoder are defined as

$$\boldsymbol{E} = [e_0, e_1, e_2, e_3]^T = \boldsymbol{F} - \boldsymbol{F}^o, \tag{9}$$
$$\boldsymbol{E}^c = [e_0^c, e_1^c, e_2^c, e_3^c]^T = \boldsymbol{F} - \boldsymbol{F}^c. \tag{10}$$

Then Øien and Lepsoy[3] showed that the relation between $\boldsymbol{E}$ and $\boldsymbol{E}^c$ can be given in equality by

$$\boldsymbol{E} = \sum_{k=0}^{\infty} \boldsymbol{L}^k \boldsymbol{E}^c = \boldsymbol{E}^c + \boldsymbol{L}\boldsymbol{E}^c + \boldsymbol{L}^2\boldsymbol{E}^c \cdots. \tag{11}$$

Most fractal coders attempt to minimize only the zero-order term $\boldsymbol{E}^c$ without considering the higher order terms, based on the collage theorem. But, in case of the CPM, more terms can also be included in the minimization without increasing the complexity of the encoder.

From the algebraic multiplications of the $\boldsymbol{L}$ matrix, we know that the $\boldsymbol{L}^4$ has a block diagonal structure. Thus by substituting the power matrices $(\boldsymbol{L}, \boldsymbol{L}^2, \boldsymbol{L}^3, \cdots)$ into Eq.(11), we can derive the following expressions of each frame's attractor error.

$$\begin{aligned} e_0 &= \sum_{k=0}^{\infty} (L_0L_3L_2L_1)^k [e_0^c + L_0e_3^c + L_0L_3e_2^c + L_0L_3L_2e_1^c] \\ &= e_0^c + \sum_{k=0}^{\infty} (L_0L_3L_2L_1)^k L_0 [e_3^c + L_3e_2^c + L_3L_2e_1^c + L_3L_2L_1e_0^c], \end{aligned} \tag{12}$$

$$\begin{aligned} e_1 &= \sum_{k=0}^{\infty} (L_1L_0L_3L_2)^k [e_1^c + L_1e_0^c + L_1L_0e_3^c + L_1L_0L_3e_2^c] \\ &= e_1^c + L_1e_0^c + \sum_{k=0}^{\infty} (L_1L_0L_3L_2)^k L_1L_0 [e_3^c + L_3e_2^c + L_3L_2e_1^c + L_3L_2L_1e_0^c], \end{aligned} \tag{13}$$

$$\begin{aligned} e_2 &= \sum_{k=0}^{\infty} (L_2L_1L_0L_3)^k [e_2^c + L_2e_1^c + L_2L_1e_0^c + L_2L_1L_0e_3^c] \\ &= e_2^c + L_2e_1^c + L_2L_1e_0^c \\ &\quad + \sum_{k=0}^{\infty} (L_2L_1L_0L_3)^k L_2L_1L_0 [e_3^c + L_3e_2^c + L_3L_2e_1^c + L_3L_2L_1e_0^c], \end{aligned} \tag{14}$$

$$e_3 = \sum_{k=0}^{\infty} (L_3L_2L_1L_0)^k [e_3^c + L_3e_2^c + L_3L_2e_1^c + L_3L_2L_1e_0^c]. \tag{15}$$

If we assume that all frame-to-frame mappings L_k $(0 \leq k < 4)$ have the same contractivity factor s, then Eqs.(12)-(15) respectively yield the inequalities by means of the triangle inequality theorem, which are given by

$$\begin{aligned}
\|e_0\| &\leq \|e_0^c\| + \sum_{k=0}^{\infty} s^{4k+1} \|e_3^c + L_3 e_2^c + L_3 L_2 e_1^c + L_3 L_2 L_1 e_0^c\| \\
&\leq \sum_{k=0}^{\infty} s^{4k} \{\|e_0^c\| + s\|e_3^c\| + s^2\|e_2^c\| + s^3\|e_1^c\|\}, \qquad (16) \\
\|e_1\| &\leq \|e_1^c + L_1 e_0^c\| + \sum_{k=0}^{\infty} s^{4k+2} \|e_3^c + L_3 e_2^c + L_3 L_2 e_1^c + L_3 L_2 L_1 e_0^c\| \\
&\leq \sum_{k=0}^{\infty} s^{4k} \{\|e_1^c\| + s\|e_0^c\| + s^2\|e_3^c\| + s^3\|e_2^c\|\}, \qquad (17) \\
\|e_2\| &\leq \|e_2^c + L_2 e_1^c + L_2 L_1 e_0^c\| + \sum_{k=0}^{\infty} s^{4k+3} \|e_3^c + L_3 e_2^c + L_3 L_2 e_1^c + L_3 L_2 L_1 e_0^c\| \\
&\leq \sum_{k=0}^{\infty} s^{4k} \{\|e_2^c\| + s\|e_1^c\| + s^2\|e_0^c\| + s^3\|e_3^c\|\}, \qquad (18) \\
\|e_3\| &\leq \sum_{k=0}^{\infty} s^{4k} \|e_3^c + L_3 e_2^c + L_3 L_2 e_1^c + L_3 L_2 L_1 e_0^c\| \\
&\leq \sum_{k=0}^{\infty} s^{4k} \{\|e_3^c\| + s\|e_2^c\| + s^2\|e_1^c\| + s^3\|e_0^c\|\}, \qquad (19)
\end{aligned}$$

Notice that each frame's attractor error is limited by two bounds. The larger bound is in terms of each frame's collage error e_k $(0 \leq k < 4)$, which is the result of direct application of the conventional collage theorem to the CPM. The smaller bound is in terms of the following 4 entities.

$$\begin{aligned}
&\bullet \quad e_0^c \\
&\bullet \quad e_1^c + L_1 e_0^c \\
&\bullet \quad e_2^c + L_2 e_1^c + L_2 L_1 e_0^c \\
&\bullet \quad e_3^c + L_3 e_2^c + L_3 L_2 e_1^c + L_3 L_2 L_1 e_0^c \qquad (20)
\end{aligned}$$

Therefore we can reduce the attractor error bound by minimizing the norm of the 4 entities in Eq.(20), instead of minimizing the norm of the collage error e_k $(0 \leq k < 4)$ independently.

By changing variables, the 4 entities in Eq.(20) can be rearranged as

$$\begin{aligned}
e_0^c &= f_0 - f_0^c \\
&= f_0 - (L_0 f_3 + T_0), \qquad (21) \\
e_1^c + L_1 e_0^c &= f_1 - f_1^c + L_1(f_0 - f_0^c) \\
&= f_1 - (L_1 f_0 + T_1) + L_1(f_0 - f_0^c) \\
&= f_1 - (L_1 f_0^c + T_1). \qquad (22)
\end{aligned}$$

Similarly,

$$\begin{aligned}
&e_2^c + L_2 e_1^c + L_2 L_1 e_0^c \\
&= f_2 - [L_2(L_1 f_0^c + T_1) + T_2], \qquad (23)
\end{aligned}$$

$$e_3^c + L_3 e_2^c + L_3 L_2 e_1^c + L_3 L_2 L_1 e_0^c$$
$$= \quad f_3 - \{L_3[L_2(L_1 f_0^c + T_1) + T_2] + T_3\}. \tag{24}$$

From these rearrangements, it can be realized that the minimization of the 4 entities in Eq.(20) leads to the new prediction method illustrated in Figure 2.(b).