Fractal Video Coding Based on Quadtree Partitioning

Mr. Bhushan U. Bawankar¹, Mr. Shailesh D. Kamble²
¹-Department of Information Technology, R.T.M.N. University, Yeshwantrao Chavan College of Engineering, Nagpur, India.
²-Department of Computer Technology, R.T.M.N. University, Yeshwantrao Chavan College of Engineering, Nagpur, India.

¹bawankar_bhushan@yahoo.co.in
²shailesh_2kin@rediffmail.com

Abstract—Fractal video compression becomes a hot topic of research due to its high compression ratio performance. However, the time required to encode the video is more, so it cannot widely used. In our proposed approach, we have used a quad-tree partitioning scheme to partition the frames of video up to that level at which no further partition is possible. With this proposed approach, our aim is to shorten the encoding time.

Keywords—Fractal video coding, quadtree, PIFS, Affine

I. INTRODUCTION

Fractal Geometry has been widely used in the fields of mathematics, physics, biology, materials chemistry, geological exploration, computer and information science. At present, the main focus of the research is towards the quick encoding for solving the problem of taking so much time in fractal encoding [1]. To solve this problem, we use quad-tree partitioning scheme with different color coding component. There are number of partitioning schemes are available such as fixed block size method, horizontal-vertical, irregular and polygonal blocks, but, some observation lead to use the Quadtree partitioning technique. In this technique, the largest range blocks are used in a quad-tree partitioning and provides best rate distortion. If the matching criterion is unsatisfied, the size of the range block is halved and the process continues until a minimum block size is reached. So, this will increase the compression ratio and maintain the clarity [2],[3],[6]. In our approach, we take a colour video as input and convert it into number of gray frames. After that, applying quadtree partitioning technique on gray frames and then use any standard fractal based approach to encode the video. Our main aim with this proposed approach is to provide fastness in video encoding to short the encoding tim and increase the compression ratio.

A. Fractal Video Coding

Fractal compression technique is based on the individual regions of an image may exhibit similarity to each other at different scales [1]. The fundamental principal was introduced by M. Barnsley in 1988 for fractal image compression is to represent an image with contractive transforms of which the fixed point is close to that image. A practical reality was given to fractal compression by Jacquin with partitioned IFS(PIFS) in 1990 [2],[3],[6]. By Jacquin, the image is divided into small image blocks i.e. sub-images with non-overlapping is called as Range blocks(R block) and find an image block which is the most similar to current R block is called as Domain block (D block) and and for each range block selecting the domain block that gives the best match to the range block.

For fractal video coding, Hurd and Gustavus[13] extended the Jaquin’s method for video compression and also used motion compensation for coding some range blocks by reporting compression ratios from 21:1 at an average PSNR of 39.2 dB to 79:1 at an average PSNR of 30.8 dB for a 160 x 120, 8-bit grayscale video sequence. Lazar and Bruton [14] extended Jacquin’s work by using 3-D blocks and reported an average compression ratio of 74.39 at an average PSNR of 32-33 dB, for the 360 x 280, 8 bit/pixel 30 Hz ‘Miss America’ video sequence. Mohammad Gharavi-Alkhansari and Thomas S. Huang[15] used matching pursuit algorithm and coded at 80 Kbits/sec with an average PSNR of 36-37 dB, 12.5 Hz ‘Miss America’ Video sequence. Yarish Bnjmohan and Stanley H. Mneney[4] used wavelet for video compression by comparing the video compression standard H.263+ on the “Akiyo” sequence and obtained average bit-rates of 0.078bpp and 0.085bpp with the average PSNR as 32.05dB and 30.82dB.

II. THE BASIS FOR FRAC TAL VIDEO CODING

A. Contractive Mapping

A transformation f : X→X on a metric space (X, d) is called contractive or contractive mappings if there exist a constant 0≤ s < 1cause d(f(x), f(y)) ≤ s.d(x, y) ∀ x, y∈ X where s is a contractive factor [1],[5],[6],[7],[8].

B. Partitioned Iterated Function System
Fractal video compression is based on Iterated Function System (IFS); the idea of IFS is to have a finite set of contraction mappings, \( \{W_n : X \rightarrow X\} \) with compression factor \( S_n \): \( n = 1, 2, ..., N \). By applying IFS to an image and provided the mapping is contractive, PIFS (Partitioned Iterated Function Systems) states that instead of finding the IFS of the entire image, it is possible to partition the image into non-overlapping blocks of ranges and then apply IFS on each block of ranges.[1],[9].

C. The Collage Theorem

In simplest case, collage is specified by an intensity map and spatial map. Let \( \{X: f_1, ..., f_n\} \) be an IFS, compression factor \( s < 1 \), \( A \) is its attractor, and \( E \in H(X) \) is any given set, there[1],

\[
h(A, E) \leq \frac{1}{1 - S} h(E, \bigcup_{i=1}^{N} f_i(E))
\]

III. COLOR CODING COMPONENT

A. Red, Green and Blue (RGB)

The human visual system is sensitive to three primary colours: red, green and blue. Therefore, these three colors are used to represent colours in digitalized images or videos. Usually 8 bits are reserved for each color, so in theory the number of colours possible equals \( 2^{24} \). In a compressing process of a color image the main idea is to divide the image into its three different layers or components (red, green and blue). It is then possible to compress each of these layers separately, therefore, the data needed to be stored and the time of encoding will be three times what it takes for the gray-scaled image[5],[7],[8],[10].

B. YIQ and YUV

There are two coordinate systems which are universally used, such as, YUV and YIQ, where Y is the luminance or brightness, I-U is the hue and Q-V is the saturation and the combination of the hue and saturation is called chrominance. These two systems related to the R, G and B by linear transformation as shown below:

\[
\begin{bmatrix}
Y \\
I \\
Q
\end{bmatrix} =
\begin{bmatrix}
0.299 & 0.587 & 0.114 \\
0.596 & -0.275 & -0.321 \\
0.212 & -0.528 & 0.311
\end{bmatrix}
\begin{bmatrix}
R \\
G \\
B
\end{bmatrix}
\]

\[
\begin{bmatrix}
Y \\
U \\
V
\end{bmatrix} =
\begin{bmatrix}
0.299 & 0.587 & 0.114 \\
-0.147 & -0.289 & 0.436 \\
0.615 & -0.515 & -0.100
\end{bmatrix}
\begin{bmatrix}
R \\
G \\
B
\end{bmatrix}
\]

There are two reasons for using YIQ and YUV is that first, it is possible to significantly reach much higher compression ratios compared with using RGB, and with only a small or negligible amount of degradation in the output image and second, the human visual system is more sensitive to the luminance than the chrominance, therefore, the I-U and Q-V can be decimated to one-half or one-quarter of their original size[5],[7],[8].

C. YCbCr

YCbCr is a family of color spaces used as a part of the color image pipeline in video and digital photography systems. Where, Y is luminance, Cb and Cr are the blue-difference and red-difference chroma components. The meaning of Y which is that light intensity is non-linearly encoded based on gamma corrected RGB primaries.

YCbCr is is a way of encoding RGB information. The actual color displayed depends on the actual RGB primaries used to display the signal. Therefore a value expressed as YCbCr is predictable only if standard RGB primary chromaticities are used.

IV. FRAC TAL VIDEO CODING USING QUADTREE

Quadtree based fractal video coding method is a way to change the size of range block and the domain block based on the ratio of allowable error \( E \), in order to achieve the purpose of getting a reasonable choice of affine transform coefficients and reduce the searching and encoding time[11],[12]. The steps are as follows

1) Segment the original image into range block of size 2x2, do not overlap each, as the initial range block.
2) Taken domain block 2 times the size of the range block in the original image, doing averaging 8z affine transformations to domain block.
3) Calculate the root of mean square of range block and each corresponding transformed domain block, as the matching error between the two blocks. If the matching error to satisfy \( d < \varepsilon, \varepsilon \) is a preset tolerable error, go to step 5.
4) If a full search completed, and did not meet the conditions \( d < \varepsilon \), segment the original block into four equal, repeat steps (2) to (4).
5) Record the current fractal coding information to the Fractal gallery, complete a fractal encoding.
Transformation used in fractal encoding is of three types. Firstly, affine transformation in which the domain block is transformed into eight small blocks as per the size of the range block and then match with the vertices of the range block. Secondly, Wavelet-Based-Fractal-Transform is based on the theory of multi-resolution analysis as well as iterated-function-system and its ability to construct the same image iteratively as well as provide local time frequency analysis on the image. At last, Discrete-Cosine-Transform form an efficient basis for image blocks due to existence of mutual orthogonality.

In dealing with gray scale images, intensity of pixels should be treated as a third spatial dimension, thus the affine transformation for the gray scale images will become:

$$w(x) = w\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} a & b & 0 \\ c & d & 0 \\ 0 & 0 & s \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + \begin{pmatrix} e \\ f \\ 0 \end{pmatrix}$$

where $s$ represents the contrast (luminance) and $o$ represents the offset to the pixel contrast which is the brightness of the transformation[2],[5],[6],[9].

VI. EVALUATION PARAMETER
The evaluation is based on the following parameter

A. Image Quality

PSNR (Peak Signal to Noise Ratio) is defined as the ratio of an original image and a coded/decoded image. PSNR is used to compute the quality of the reconstructed image compared with the original image by measuring the differences between the two images; the formula used to compute the PSNR is,

$$\text{PSNR} = 20 \log_{10} \left( \frac{b}{\text{rms}} \right)$$

where $b$ is the highest pixel value (255) and rms is the root mean square differences between the two images. PSNR is measured in dB, the highest the PSNR, the better the quality of the reconstructed image [7].

B. Compression Ratio

Compression ratio is defined as the ratio of an original image and compressed image.

$$\text{Compression Ratio} = \frac{\text{Org. Img. Size}}{\text{Comp. Img. Size}}$$
VII. EXPERIMENTAL ANALYSIS

The process realized in MATLAB environment. In this, the range block is taken of size 4 X 4 and the domain block size is double of range block that is, 8 X 8. In this, we take one video as input and convert it into number of frames and for testing we apply the quadtree method to the single frame and the result is as follows with original and compressed image. For testing, we use standard traffic video as a test video.

(1) Select Traffic video as test video

![Original Image](image1)

![Partitioned Image](image2)

![Compressed Image](image3)

VIII. CONCLUSION

This paper briefly introduces the basis related to the fractal and some colour coding component and discusses the fractal coding steps based on quadtree. Experimental result shows the difference between the original image and compressed image. Our proposed approach can shorten the encoding time and improve the compression ratio and image quality.

REFERENCES


