Efficient Fractal Image Coding using Fast Fourier Transform

S.B Dhok, R.B. Deshmukh, A.G. Keskar
Visvesvaraya National Institute of Technology, Nagpur (INDIA)
sanjaydhok@gmail.com, rbdeshmukh@vnit.ac.in, agkeskar@vnit.ac.in

Abstract - The fractal coding is a novel technique for image compression. Though the technique has many attractive features, the large encoding time makes it unsuitable for real time applications. In this paper, an efficient algorithm for fractal encoding which operates on entire domain image instead of overlapping domain blocks is presented. The algorithm drastically reduces the encoding time as compared to classical full search method. The reduction in encoding time is mainly due to use of modified crosscorrelation based similarity measure. The implemented algorithm employs exhaustive search of domain blocks and their isometry transformations to investigate their similarity with every range block. The application of Fast Fourier Transform in similarity measure calculation speeds up the encoding process. The proposed eight isometry transformations of a domain block exploit the properties of Discrete Fourier Transform to minimize the number of Fast Fourier Transform calculations. The experimental studies on the proposed algorithm demonstrate that the encoding time is reduced drastically with average speedup factor of 538 with respect to the classical full search method with comparable values of Peak Signal To Noise Ratio.

Index Terms - similarity measure, fast fourier transform, fractal coding, fractal image compression.

I. INTRODUCTION

Fractal image coding is a promising compression technique which was pioneered by Michael Barnsley and A.E. Jacquin in 1990. The technique is theoretically capable of achieving very high compression ratios. The first practical implementation scheme of fractal encoding algorithm was based on partitioned iterated function system (PIFS) which exploits image redundancy for coding through block self-similarity[1,2]. The efficiency of image compression method depends on mean square error (MSE) between original and reconstructed image, compression ratio and time required for encoding and decoding. Fractal image coding provides high reconstructed image quality, resolution independence and fast decoding[3]. The major drawback of fractal based compression is large amount of time required for encoding the images. Very high encoding time in classical exhaustive search fractal compression method [4] is due to large number of comparisons between every range block in the range image and all possible overlapping domain blocks in the domain image. This involves RMS distance calculations between range and domain blocks. The frequent accessing of non-overlapping domain blocks instead of entire image also slows down the search process. The other bottleneck in the encoding algorithm is computationally intensive equations for calculating contrast and brightness parameters of the fractal transformations. In order to increase the domain pool for finding the better match, normally eight isometric transformations of every domain block is considered. This also increases the search time considerably.

Many strategies are used for reducing encoding time. The reduction in encoding time in currently available fractal methods is largely due to limited domain search based on classification of domain blocks[5]. Instead of searching for matching domain block in entire pool of possible domain blocks, the search is limited to a class of domain blocks having approximately similar average intensity and variance or having similar structures like peripheral or central blocks[4]. The search time can also be reduced by using genetic algorithm or other evolutionary search algorithm like particle swarm optimization [6]. There is usually very high correlation amongst adjacent range blocks such that the domain blocks mapped by the previous adjacent range blocks can be considered as good candidate blocks for priority search[7].

In order to increase the probability of finding better match, normally eight isometries of domain block obtained by rotation and reflection are considered. This further increases the encoding time. A fast algorithm can remove all of the redundant computations for the isometries by performing MSE computations in DCT domain[8]. Furthermore since most of the energy of the data is concentrated in low frequencies, MSE computations can be quickly done using only low frequency data[9]. A fast algorithm based on single kick out condition and zero contrast prediction can avoid large number of range – domain block matches[3]. Encoding complexity can also be reduced by domain classification method based on approximation error measure and deferring the range-domain comparisons with respect to preset block[10]. Domain block searching efficiency can be enhanced by employing support vector machines trained to recognize self similar pattern features[11]. Other fast encoding techniques available in literature are based on IFS with probabilities [12], using predefined values for contrast scaling parameters which are estimated through entropy measure [13], prediction and subblock based algorithm[14] or employing Shannon entropy based classification[15]. The multiresolution analysis of wavelet theory is also tried for fractal encoding [16].
Comparatively few methods are available for speeding up the exhaustive search methods. M. Ramkumar and G.V.Anand proposed Fast Fourier Transform (FFT) based method to speed up the process of matching domain and range blocks. In this method isometric transformations of domain block are chosen to permit block matching through crosscorrelation of range and domain blocks which is implemented through FFT. Another method for lossless acceleration of exhaustive search method is proposed by D.Saupé [18,19]. In this method the MSE, contrast and brightness parameters are expressed as inner product computations. These inner products are expressed as the finite impulse response of the downfiltered image with respect to range. The filtering operation is done in frequency domain via FFT to speedup the computations. In both these methods the domain image is downsampled by averaging the adjacent pixels before applying FFT. This halves the number of domain blocks to be searched. The piece-wise self-similarity is more evident in the fractal coding if large number of domain blocks are used [20]. Both these methods essentially use L2 metric to calculate all the MSEs, contrast and brightness parameters before discarding the block.

In this paper, we implement an efficient fractal coding method in which maximum domain pool is utilized by upsampling the range block. In our proposed method, we use modified similarity measure based on normalized mean subtracted cross-correlation to determine the similarity between range and domain blocks. The similarity measure is well suited for FFT based implementation. This approach also eliminates lots of redundant contrast and brightness parameter calculations due to ranking of domain blocks by similarity index. We also use in this paper eight isometry transformations for a domain block such that their FFTs can be easily obtained from the FFT of the domain block by exploiting the properties of two dimensional Discrete Fourier Transform.

II. REVIEW OF FRACTAL IMAGE CODING

Fractal image compression is based on theory of partitioned iterated function systems. Let X be a complete metric space with metric d, A ∈ X be any image and w : X → X be a contractive transformation. Then the repeated application of transformation w on A will take it to the fixed point of transformation Â ∈ X such that w(Â) = Â. The collage theorem states that if d(A, w(A)) is small then the distance d(Â, A) is also small [4].

Usually w is of the form

\[ w = U_{w_i}, \quad i = 1, 2, \ldots, P \]  

where each w_i is a contractive mapping and P denotes number of partitions. Each w_i mapping consists of a spatial mapping and a grayscale mapping. Associated with each w_i is a range block R_i and a domain block D_i. The range blocks which are generally of equal size, are non-overlapping and tile the image completely. The domain blocks are overlapping blocks in the image. The affine transformations are chosen such that they map appropriate D_i into R_i. The size of each D_i is larger than that of the corresponding R_i to ensure contractivity. Usually the size of each D_i is twice the size of each R_i. The fractal transformation can be represented by

\[ R_i = s_i T_i \{ Q(D_i) \} + o_i \]  

where s_i is the scale parameter, o_i is the brightness parameter, Q is a contractivity operator and T_i […] is an isometry transformation for partition i. The range of allowable values of s_i and o_i are constrained by convergence conditions as applicable to contractive PIFS transformation. The isometry transformation T_i […] consists of spatial rotation and/or inversion of domain block. Usually eight isometry transformations of every domain block are used to increase the search pool of domain blocks. The goal of fractal transformation is to search for every R_i the best matching domain block D_i from available pool of all the domain blocks such that R_i is approximately equal to R_i as given by (2).

The decoded image is obtained by applying the fractal transformation (2) on an arbitrary initial image and iterating the process on transformed image number of times. The decoded image is guaranteed to be ‘close’ to the original image as per Collage theorem. In practice, for images, convergence is achieved in 8 to 10 iterations. The ‘closeness’ of the reconstructed image to the original image is usually represented by the Peak Signal to Noise Ratio (PSNR) of the decoded image.

III. FFT BASED SIMILARITY MEASURE

Let us assume that initially range image H* and domain image X* are same and of size N x N pixels. The range image H* is partitioned in non-overlapping blocks of size N x N. Let H be a range block of size N x N with mean graylevel of H. The domain image X* is partitioned into overlapping blocks X_{ij} of size 2N x 2N and then the domain block is downsampled by factor of two in both the dimensions. Let X_{ij} be the downsampled domain block of size N x N with mean graylevel of X_{ij}, where i and j are x and y coordinates of start pixel of domain block respectively. Let T_i[X_{ij}] represents isometry transformed domain block where k = 0, 1, 2, ..., 7.

The degree of similarity between range block H and domain block X_{ij} is given by normalized mean subtracted cross-correlation coefficient C_{d(i,j,k)} as

\[ C_{d(i,j,k)} = \langle H_1, T_i[X_{ij}] \rangle / \| H_1 \| \| X_{ij} \| \]  

where H_1 = mean subtracted range block

\[ X_{ij} = \text{mean subtracted domain block} \]

\[ X_{ij} = X_{ij} - X_{0j} \]

In (3), operator \( \langle, \rangle \) represents inner product operation and \( \| . \| \) represents L2 norm. The MSE will be minimised by maximizing the absolute value of coefficient C_{d(i,j,k)} [17,21].

Since H_1 is zero mean block, it can be easily proved that

\[ \langle H_1, T_k[X_{ij}] \rangle = \langle H_1, T_k[X_{0j}] \rangle \]  

Also instead of starting with the larger domain block X_{2ij} of size 2N x 2N and then decimating it to get N x N sized domain block X_{ij}, we can keep larger domain block X_{2ij} unchanged in the domain image and consider zero
interlaced larger mean subtracted range block $H_2$ of size $2N_1x2N_1$ where,

$$H_{2}(m,n) = H_1(m/2,n/2) \text{ if } m,n = \text{ even}$$

$$= 0 \text{ if } m,n = \text{ odd} \quad (5)$$

$$X_{ij}(m,n) = X_{ij}(2m,2n) \quad (6)$$

It can be seen that

$$<H_1, T_k[X_{ij}]> = <H_2, T_k[X_{0ij}]> \quad (7)$$

To get inner product of $H_1$ with next domain block, we just need to translate the $H_2$ block by suitable $i$ and $j$ shift and repeat the process. This is similar to cross-correlation operation. The modified similarity measure can be expressed as

$$C(i,j,k) = \langle H_2, T_k[X_{0ij}] \rangle / \langle H_1 \parallel X_{ij} \parallel \rangle \quad (8)$$

Let us define the similarity index matrix as

$$C = \{C(i,j,k) \text{ for all } i,j,k \} \quad (9)$$

The similarity index matrix gives the degree of similarity between a range block and all the possible domain blocks for all the eight possible isometric transformations.

### IV. EFFICIENT FRACTAL ENCODING ALGORITHM

The goal of fractal encoding is to find for every range block $H_1$, a certain domain block $X_{0ij}$ with associated isometry transformation $T_{k}[\cdot]$ such that affine transformation of $X_{0ij}$ closely matches $H_1$ i.e.

$$H = s \cdot T_{k}[X_{0ij}] + o \quad (10)$$

Where $s,o$ are contrast and brightness parameters. The transformation is coded in terms of $s$, $o$, $i$, $j$ and $k$ parameters.

Thus

$$H_1 = s \cdot T_{k}[X_{1ij}] \quad (11)$$

$$H_0 = sX_{0ij} + o \quad (12)$$

Now if we use folded $H_2$ block as given by (5), all the cross-correlation coefficients between $H_1$ and all the domain blocks can be calculated at once by linear convolution of folded $H_2$ with domain image $X^*$. Using properties of Discrete Fourier Transform (DFT), the linear convolution of two sequences can also be expressed as circular convolution of suitably zero padded sequences. Hence these cross-correlation coefficients can also be calculated by circular convolution of suitably zero-padded & folded $H_2$ and zeropadded domain image $X^*$ via two dimensional FFT algorithm. This will speed up the calculations. The advantage of above method in which domain image is of same size as range image but domain block considered is larger, will be availability of larger pool of domain blocks compared to the method in which domain image is decimated by factor of two in both $x,y$ directions. Another advantage of this method is that FFT of $X^*$ has to be calculated only once and can be repeatedly used for all the range blocks.

The calculation of norm $\parallel H_1 \parallel$ is straightforward but calculation of norm $\parallel X_{0ij} \parallel$ is not very easy since $X_{0ij}$ represents mean subtracted domain block where every domain block has different mean. The calculations can be simplified by using the following relation,

$$\parallel X_{0ij} \parallel = [(\sum X_{0ij}^2 - N_{ij}^2 X_{0ij}^2)]^{1/2} \quad (13)$$

The squared sum matrix $\sum X_{0ij}^2$ of all the domain blocks can be calculated at once by convolving mask $M_2$ with $X^*$ containing pixel by pixel squared values of entire domain image $X^*$ where $M_0 = \text{matrix of size } N_2xN_2$ containing all 1’s and $M_1$ is matrix of size $2N_1x2N_1$ obtained by zero interlacing of $M_0$.

$$M_i(m,n) = M_i(m/2,n/2) \text{ if } m,n = \text{ even}$$

$$= 0 \text{ if } m,n = \text{ odd} \quad (14)$$

The mean matrix of all the domain blocks (all $X_{0ij}$’s) can be calculated at once by convolving mask $M_2$ with entire domain image $X^*$ where $M_0 = \text{matrix of size } N_2xN_2$ where $M_1$ is given by (14). Fast convolutions are done via FFT similar to cross-correlation calculation.

$$\text{Norm} \parallel X_{0ij} \parallel \text{ calculation, though involved, needs to be done only once and is used for all the range blocks. Finally similarity index matrix can be calculated. Practically for most similar domain block absolute values of } C(i,j,k) \text{ will be close to 1. The contrast parameter ‘s’ and brightness parameter ‘o’ can be shown to be given by}$$

$$s = \langle \parallel H_1 \parallel^2 \rangle / \langle \langle H_2, T_{k}[X_{2ij}]\rangle \rangle \quad (15)$$

$$o = H_0 - sX_{0ij} \quad (16)$$

The maximum value of similarity index satisfying the constraints on the contrast and brightness parameters is selected for encoding. The position of this maximum value gives the $x$-coordinate and $y$-coordinate of the best fitting domain block.

In the proposed method, exhaustive search for most similar domain block is carried out. Instead of calculating similarity index in spatial domain for each range and domain block, the properties of Fourier Transform are exploited to calculate similarity index between a range block and all possible domain blocks at once in frequency domain via FFT. This involves operation on entire domain image instead of overlapping domain blocks. The contrast and brightness parameters of fractal coding are estimated using computationally less intensive equations. In fact these parameters are calculated using the partial results during the course of calculation of similarity index matrix. This reduces the encoding time.

Instead of applying isometry transformations $T_{k}[\cdot]$ on domain block, the complementary isometry transformations $T_{k'}[\cdot]$ are applied on range block during fractal image encoding. During fractal image decoding, isometry transformations $T_{k}[\cdot]$ are applied on domain block. Since our method works on entire domain image instead of individual domain blocks, complementary isometry transformations can be conveniently applied to range block with similar results. These transformations are defined in Table I. Since these complementary isometry transformed range blocks are spatially related to original range block through suitable rotations and/or inversions, their FFTs can be expressed in terms of FFT of original range block. These FFTs thus need not be calculated separately but can be obtained using properties of two dimensional Discrete
Fourier Transform. This also helps in increasing the search speed of the algorithm.

**TABLE I**

<table>
<thead>
<tr>
<th>k</th>
<th>T_k[.].</th>
<th>T_k^c[.].</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>No change</td>
<td>No change</td>
</tr>
<tr>
<td>1</td>
<td>Inversion in x-direction</td>
<td>Inversion in x-direction</td>
</tr>
<tr>
<td>2</td>
<td>Inversion in y-direction</td>
<td>Inversion in y-direction</td>
</tr>
<tr>
<td>3</td>
<td>Inversion in both x and y-direction</td>
<td>Inversion in both x and y-direction</td>
</tr>
<tr>
<td>4</td>
<td>Transpose</td>
<td>Transpose</td>
</tr>
<tr>
<td>5</td>
<td>Inversion in x-direction followed by transpose</td>
<td>Transpose followed by inversion in x-direction</td>
</tr>
<tr>
<td>6</td>
<td>Inversion in y-direction followed by transpose</td>
<td>Transpose followed by inversion in y-direction</td>
</tr>
<tr>
<td>7</td>
<td>Inversion in both x and y-direction followed by transpose</td>
<td>Transpose followed by inversion in both x and y-direction</td>
</tr>
</tbody>
</table>

V. SIMULATION RESULTS

The proposed algorithm is simulated in Matlab7 on Intel Core Duo CPU at 2.66 GHz with 2 GB RAM. The experimentation is done on number of test images of “lena”, “pepper”, “cameraman”, “fishingboat”, “baboon” etc for different values of image sizes and block sizes.

The ‘s’ and ‘o’ parameters are quantized using 5 and 7 bits respectively. The isometry transformation parameter ‘k’ is represented by 3 bits. To ensure the convergence the |s| is normally chosen less than 1. But larger values of ‘s’ also give eventual convergence for PIFS [22]. Based on the experimentation, the range of ‘s’ and ‘o’ is restricted to \( -1.8 < s < 1.8 \) and \( -256 < o < 256 \) respectively for contractive mapping constraint. The corresponding s, o, i, j and k are stored as parameters of fractal transform of range block H. The procedure is repeated for all the range blocks.

The fractal decoding process is quick and straightforward. We start with any initial image as domain image \( X_d^* \) of same size as the original image which is encoded. We also start with another blank image as range image \( X_d^* \) of same size as the original image. We call this image as the reconstructed image. We partition this reconstructed image in nonoverlapping range blocks of size \( N_x \times N_y \), as we have done in encoding process. We take transform parameters of first range block s,o,i,j and identify the domain block of size \( 2N_x \times 2N_y \) starting with x & y coordinates of i,j in the initial image \( X_d^* \). We downsample this domain block by factor of 2 in both x & y direction to get reduced domain block \( X_d^* \) of size \( N_x \times N_y \). The isometry transformations \( T_k[.]. \) is applied to domain block \( X_d^* \). The graylevel transformation is applied to this transformed domain block to get the corresponding range block \( H_d \), i.e. \( H_d = sT_k[X_d^*] + o \). This decoded range block is pasted in its correct partition position in the reconstructed image \( H_d^* \). The process is repeated for all the range blocks to get complete reconstructed image \( H_d^* \) after first iteration. In next iteration we take \( X_d^* = H_d^* \) and repeat the entire procedure. The reconstructed image \( H_d^* \) usually converges to original image \( H^* \) after 5 to 10 iterations.

The compression ratio (CR) is defined as the ratio of number of bits required for original image to the number of bits required for encoded image.

For implemented method, CR parameter is given by

\[
CR = \{N_xN_y + N_k + N_x + N_y\} \quad (17)
\]

where \( N_k \) is the number of bits per pixel used in original image and \( N_x, N_y, N_k, N_x, N_y \) denote number of bits used for representation of fractal parameters of i,j,k,s and o respectively.

The MSE between original and decoded image is given by

\[
MSE = \{\sum_{p,q} [H^*(p,q) - H_d^*(p,q)]^2\} / \{N_xN_y\} \quad (18)
\]

where \( H^* \) and \( H_d^* \) represents original and final decoded image of size \( N_xN_y \).

The PSNR of decoded image (in dB) is given by

\[
PSNR = 10 \log_{10} (255^2 / MSE) \quad (19)
\]

Sample values of these performance parameters for some of the test images for implemented algorithm are given in Table II. The comparative results for encoding time and PSNR of the proposed method and full search method is shown in Table III. The reconstructed “fishingboat” image of size 512x512 with blocksize of 8x8 is shown in Fig. 1. The rate distortion curves of PSNR in dB versus CR for proposed method with isometric transformations and without isometric transformations are given in Fig. 2.

**TABLE II**

<table>
<thead>
<tr>
<th>Image</th>
<th>Compression Ratio</th>
<th>Encode time (s)</th>
<th>PSNR (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lena</td>
<td>4.13</td>
<td>175</td>
<td>31.77</td>
</tr>
<tr>
<td></td>
<td>9.29</td>
<td>49</td>
<td>26.61</td>
</tr>
<tr>
<td></td>
<td>16.52</td>
<td>45.19</td>
<td>24.38</td>
</tr>
<tr>
<td></td>
<td>25.81</td>
<td>31.28</td>
<td>22.94</td>
</tr>
<tr>
<td>Pepper</td>
<td>37.16</td>
<td>20.73</td>
<td>21.31</td>
</tr>
<tr>
<td></td>
<td>4.13</td>
<td>174.53</td>
<td>36.12</td>
</tr>
<tr>
<td></td>
<td>9.29</td>
<td>48.31</td>
<td>31.71</td>
</tr>
<tr>
<td></td>
<td>16.52</td>
<td>44.69</td>
<td>29.07</td>
</tr>
<tr>
<td>Camera-man</td>
<td>25.81</td>
<td>30.37</td>
<td>27.18</td>
</tr>
<tr>
<td></td>
<td>37.16</td>
<td>20.57</td>
<td>25.70</td>
</tr>
<tr>
<td>Fishing-boat</td>
<td>4.13</td>
<td>176.19</td>
<td>29.51</td>
</tr>
<tr>
<td></td>
<td>9.29</td>
<td>49.72</td>
<td>24.85</td>
</tr>
<tr>
<td></td>
<td>16.52</td>
<td>47.26</td>
<td>21.83</td>
</tr>
<tr>
<td></td>
<td>25.81</td>
<td>30.96</td>
<td>18.97</td>
</tr>
<tr>
<td></td>
<td>37.16</td>
<td>22.04</td>
<td>18.23</td>
</tr>
</tbody>
</table>

The PSNR of decoded image (in dB) is given by

\[
PSNR = 10 \log_{10} (255^2 / MSE) \quad (19)
\]

©2011 GSTF
VI. CONCLUSION

It is found that encoding time is reduced drastically in the implemented method with equivalent PSNR values as compared to classical exhaustive search method. The average speedup factor of about 538 is obtained for the proposed method. The speedup factor increases as the size of image increases. The FFT based similarity measure is useful for decreasing the encoding time of fractal image compression. The FFT compatible isometry transformation also reduces the number of FFT computations considerably. The use of isometry transformation improves the average value of PSNR of decoded image by about 2.5 dB. The perceptual quality of decoded image is comparable to full search method with negligible blocking effects since similarity measure peaks for highly correlated domain blocks rather than minimizing the MSE.

The FFT based similarity measure can be applied to limited domain search methods with suitable modification to further speed up the encoding process. The proposed algorithm can be combined with quadtree partitioning technique to improve the quality of decoded image. The new isometry transformations which are compatible with FFT based similarity measure can be proposed for finding best matched domain block.

REFERENCES


TABLE III

COMPARATIVE RESULTS FOR PROPOSED METHOD AND FULL SEARCH METHOD

<table>
<thead>
<tr>
<th>(Image Size)/ (Block Size)</th>
<th>Proposed Method</th>
<th>Full Search Method</th>
<th>Speed-up factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>[(64x64)/(4x4)]</td>
<td>0.8</td>
<td>23.25</td>
<td>257</td>
</tr>
<tr>
<td>[(100x100)/(5x5)]</td>
<td>2.31</td>
<td>22.55</td>
<td>177</td>
</tr>
<tr>
<td>[(128x128)/(5x5)]</td>
<td>3.96</td>
<td>20.38</td>
<td>1470</td>
</tr>
<tr>
<td>[(256x256)/(16x16)]</td>
<td>15.24</td>
<td>20.11</td>
<td>14530</td>
</tr>
</tbody>
</table>

FIGURE 1

RECONSTRUCTED “FISHINGBOAT” IMAGE OF SIZE 512X512 WITH BLOCK SIZE OF 8 X 8

FIGURE 2

RATE-DISTORTION CURVES FOR PROPOSED METHOD


S.B.Dhok is Assistant Professor in Electronics Engg. Dept at Visvesvaraya National Institute of Technology, Nagpur. He has published several papers in national and international conferences. His area of interest includes Digital Image Processing, Data Compression, Digital Signal Processing. Email: sbdhok@vnit.ac.in, sanjaydhok@gmail.com

R.B. Deshmukh is Associate Professor in Electronics Engg. Dept at Visvesvaraya National Institute of Technology, Nagpur. He is a member of IEEE. He has published several papers in International Journal and Conferences. His area of interest includes Digital Design, CMOS Design, Reconfigurable FPGA Design, VLSI Signal Processing. Email: rbdeshmukh@vnit.ac.in

A.G. Keskar is Professor in Electronics Engg. Dept at Visvesvaraya National Institute of Technology, Nagpur. He is a senior member of IEEE. He has published 18 Journal papers and published 41 papers in national and international conferences. His area of interest includes Fuzzy logic, Embedded System and Machine Vision. Email: agkeskar@vnit.ac.in