A Phase Span Search Method for Fractal Image Coding

Ming-Sheng Wu
Dept. of Electrical Engineering, Cheng Shiu University, Kaohsiung, Taiwan
E-mail: sheng@csu.edu.tw

Abstract

In this paper, a phase span search method on the frequency domain is proposed to speedup the fractal encoder. The method searches for the best matched solution in the frequency domain. A coordinate system is constructed using the two lowest discrete cosine transformation (DCT) coefficients of image blocks. By mapping image blocks into the coordinate system, some image blocks with similar edge shapes will be concentrated on a specific region. Therefore the purpose of speedup can be reached by limiting the search space on the specific region. Experimental results show that, under the penalty of decaying 1 dB, the encoding speed of the proposed method is about 50 times faster than that of the full search method.

Keywords: Fractal image compression; Frequency domain; Neighborhood region search method.

1. Introduction

Fractal image compression was first proposed in 1985 by Barnsley originated from Iterated Function System (IFS) [7]. The practical coding algorithm has not realized until 1992 by Jacquin [1]. The underlying idea of the coding scheme is based on the Partitioned Iteration Function System (PIFS) which utilized the self-similarity characteristic in a nature image to achieve the purpose of compression [4, 5].

The encoding process of the fractal image compression is very time-consuming. The reason is that most of the encoding time is spent on a large amount of computations of similarity measure in order to find the best matched domain block for each range block. Hence the main research direction for fractal image compression is focused on how to reduce the encoding time. Many encoding techniques were presented by the researchers to speedup the fractal encoder. These techniques include classification techniques [3, 10, 14], quad-tree technique [6, 9, 12], and evolutionary computation technique [2, 13] etc.. Besides, the idea of the spatial correlation was also added in the fractal encoding algorithm to reduce the search space. Truong et al. [11] limit the search space for the current range block on the neighborhoods of the matched domain blocks of the neighboring range blocks by utilizing the spatial correlations between neighboring blocks in both the domain pool and the range pool. In comparison to the full search method, the algorithm achieves 2.6 times speedup ratio and obtains good retrieved image quality. Wu et al. [8] propose a spatial correlation genetic algorithm (SC-GA) to speed up the encoder. By the spirit of spatial correlations stated in [11], the SC-GA method for the current range block executes GA evolutionary process on the neighborhoods of the matched domain blocks of the neighboring range blocks. Experimental result shows that the encoding time of the SC-GA method is 1.5 times faster than that of traditional GA method under the premise of maintaining same retrieved image quality.

In this paper, a phase span search method on the frequency domain is proposed to speedup the fractal encoder. The method executes the optimal search process in the frequency domain. A coordinate system is constructed from two discrete cosine transformation (DCT) coefficients of image blocks: the lowest vertical coefficient and the lowest horizontal coefficient. The reason of executing the optimal search process in the frequency domain is that, by mapping all the range and domain blocks into the coordinate, those blocks with similar edge shapes will concentrate together. Hence we can search for the best matched domain block in the specific region including the range block. Experimental results show that, under the penalty of 1dB decay of the retrieved image quality, the encoding speed is about 50 times faster than that of the full search method.

The rest of the paper is organized as follows. We introduce the conventional fractal image coding scheme in Section 2. Section 3 describes the phase span search method on the frequency domain. Section 4 shows some experimental results to verify the performances of the proposed method. Finally, a conclusion is made in Section 5.

2. Fractal Image Encoding

The fractal image compression is based on the local self-similarity property in a nature image. The fundamental idea is coming from the Partitioned Iteration Function System (PIFS). Suppose the original gray level image \( f \) is of size \( 256 \times 256 \). Let the range pool \( R \) be defined as the set of all non-overlapping blocks of size \( 8 \times 8 \) of the image \( f \).
which makes up \((256/8)^2 = 1024\) blocks. For obeying the Contractive Mapping Fixed-Point Theorem, the domain block must exceed 2 times than the range block in length. Thus, let the domain pool \(D\) be defined as the set of all possible blocks of size \(16 \times 16\) of the image \(f\), which makes up \((256-16+1)^2 = 58081\) blocks. For each range block \(v\) from the \(R\), the fractal affine transformation is constructed by searching all of the domain blocks in the \(D\) to find the most similar one and the parameters representing the fractal affine transformation will form the fractal compression code of \(v\).

To execute the similarity measure between range block and domain block, the size of the domain block must be first sub-sampled to \(8 \times 8\) such that its size is the same as \(v\). Let \(u\) denote a sub-sampled domain block. The similarity of two image blocks \(u\) and \(v\) of size \(n \times n\) is measured by mean square error (MSE) defined as

\[
MSE(u, v) = \frac{1}{n \times n} \sum_{j=0}^{n-1} \sum_{i=0}^{n-1} (u(i, j) - v(i, j))^2
\]

(1)

![Figure 1. The diagram of eight transformations in the Dihedral group.](image)

The fractal affine transformation allows the eight transformations of the domain block \(u\) in the Dihedral. The eight transformations \(T_k : k = 0, 1, \ldots, 7\) can be expressed by the diagrams in Fig. 1, in which the origin of \(u\) is assumed to locate at the center of the block. By the eight transformation, eight transformed blocks are generated and denoted by \(u_k : k = 0, 1, \ldots, 7\), respectively, where \(u_0\) is equal to the original sub-sampled domain block \(u\). \(T_0\) picks the origin block \(u\). \(T_1\) and \(T_2\) are the flip of \(u\) with respect to horizontal and vertical lines, respectively. \(T_3\) is the flip of \(u\) with respect to both horizontal and vertical lines. \(T_4\), \(T_5\), \(T_6\), and \(T_7\) are the transformations which flip the \(u_0\), \(u_1\), \(u_2\), and \(u_3\) along the main diagonal line \(y = x\), respectively. Thus for a given block from the range pool, there are \(58081 \times 8 = 464,648\) MSE computations must be done in order to obtain the most similar block from the domain pool. Thus, in total, one needs \(1024 \times 464,648 = 475,799,552\) MSE computations to encode the whole image using this full search compression method.

The fractal affine transformation also allows the contrast scaling \(p\) and the brightness offset \(q\) on the transformed blocks. Thus the similarity is to minimize the quantity \(d = \| p \cdot u_k + q - v \|\). Here, \(p\) and \(q\) can be computed directly by

\[
p = \frac{[N < u_k, v > - < u_k, \bar{v} >, \bar{v} >]}{[N < u_k, u_k > - < u_k, \bar{v} >]} \quad \text{and} \quad q = \frac{1}{N} [v, \bar{v} - p u_k, \bar{v} ]
\]

, respectively, where \(N\) is the number of pixels of the range block and \(\bar{v} = [1 \ 1 \ \ldots \ 1]^T\).

Finally, as \(u\) runs over all the 58081 blocks in the domain pool, a set of parameters \(t_x, t_y, p, q, k\) are obtained and constitute the fractal compression code of \(v\), in which \(t_x\) and \(t_y\) represent the position of the domain block. For \(256 \times 256\) image, both \(t_x\) and \(t_y\) require 8 and 8 bits, respectively. For contrast \(p\), brightness \(q\), and the Dihedral transformation \(k\), 5, 7 and 3 bits are required, respectively. Hence one needs 31 bits in total to encode a range block. Finally, as \(v\) runs over all 1024 blocks in the range pool, the encoding process is completed.

3. The Fast Search Method on the Frequency Domain

In this section, we propose a phase span search method on frequency domain. In the proposed method, a 2-dimentsional coordinate system of frequency domain is built and all the range blocks and domain blocks in the image is mapped into the coordinate system. The system is built according to the lowest frequency DCT coefficients so that blocks of similar edge properties will scatter together. Hence, a good matched solution can be found, only if we limit the search space on the specific region including the range block.

The coordinate system of frequency domain is set up by using two DCT coefficients: the lowest vertical coefficient \(F(1, 0)\) and the lowest horizontal coefficient \(F(0, 1)\). The quantity \(F(m, n)\) is the DCT of an image block \(f(i, j)\) of size \(N \times N\) defined by

\[
F(m, n) = \frac{2}{N} C_m C_n \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} f(i, j) \cos \left( \frac{(2i+1)m\pi}{2N} \right) \cos \left( \frac{(2j+1)n\pi}{2N} \right)
\]

(2)
where \( m, n = 0, 1, \cdots, N - 1 \) and
\[
C_k = \begin{cases} 
\sqrt{2}, & \text{if } k = 0 \\
1, & \text{otherwise}
\end{cases}
\]
Typically, for \( N = 8 \), we have
\[
F(1, 0) = \frac{\sqrt{2}}{8} \sum_{i=0}^{7} \sum_{j=0}^{7} f(i, j) \cos \theta_i \quad \text{and} \quad (3)
\]
\[
F(0, 1) = \frac{\sqrt{2}}{8} \sum_{i=0}^{7} \sum_{j=0}^{7} f(i, j) \cos \theta_j ,
\]
where \( \theta_i = (2i+1)\pi/16 \), \( i = 0, 1, \cdots, 7 \). The magnitude of \( F(1, 0) \) reflects the intensity variation between the left half and right half of the image block \( f \) and the magnitude of \( F(0, 1) \) reflects the intensity variation between the upper half and lower half. Rough edge shapes of image blocks can be associated to these two coefficients.

For each image block, we calculate the coefficients \( F(1, 0) \) and \( F(0, 1) \). In terms of Fractal image compression due to the Dihedral transformation as stated above, we take the absolute values of the two coefficients to constitute a pair \((|F(1, 0)|, |F(0, 1)|)\) which represent the image block. In this study, all the range blocks and domain blocks are mapped into the coordinate system in a way that we associate an image block \( f \) to the pair \((|F(1, 0)|, |F(0, 1)|)\) in the coordinate. Assume the element with maximal norm is \(|F(1, 0)|, |F(0, 1)|\). To facilitate later search strategy, we scale all the elements with respect to this element and denote the new elements as \((|F(1, 0)|, |F(0, 1)|)\). The illustration is shown in Fig. 2 in which \((|F(1, 0)|, |F(0, 1)|)\) is the scaled element of the one with maximal norm. Thus, all elements are further mapped into the unit circle in the first quadrant.

Since the DCT operation is linear, for any two blocks \( u \) and \( v \) such that \( u = \alpha v \) for some constant \( \alpha \), then \( u \) and \( v \) must locate on the same ray from the origin. In other words, these two blocks and the blocks near this ray will have similar edge features. Therefore, for each range block, a good candidate match will exist in the neighboring region of this ray.

Based on the argument above, a search strategy of phase span is presented to reduce the search space. For each range block \( v_j \), the line \( \phi = \theta_j \) is calculated by \( \theta_j = \tan^{-1}(|F(0, 1)|/|F(1, 0)|) \) where \(|F(1, 0)| > 0 \). Then, a search space \( \Phi_j \) is built by taking \( \phi = \theta_j \) as the center and spanning a phase angle \( \Delta \theta \) toward two sides. In addition, one takes the line \( \phi = 45^\circ \) as the symmetry axis.

The region \( \Phi_j \) reflected from \( \Phi_i \) by \( \phi = 45^\circ \) is also a candidate region as the Dihedral transformations are considered. These two regions depicted in Fig. 3 are the restricted search space for \( v_j \).

![Figure 2. The distribution condition of image blocks on frequency domain.](image)

![Figure 3. The search space of phase angle span search method.](image)

The detailed steps for phase span search method are stated as follows:

1. For all of the range and domain blocks, calculate the \( F(1, 0) \) and \( F(0, 1) \) from (3) and (4), respectively, and constitute their DCT absolute coefficient pair \((|F(1, 0)|, |F(0, 1)|)\).
2. Find the scaled DCT absolute coefficient pair \((|F(1, 0)|, |F(0, 1)|)\).
3. For each range block \( v_j \), perform steps 4-6.
4. Find the straight line \( \phi = \theta_j \).
5. Take the two straight lines: \( \phi = \theta_j \) and \( \phi = 90^\circ - \theta_j \) as the center lines. Span a phase...
angle \( \Delta \theta \) toward their two sides, respectively, to constitute the search spaces \( \Phi_1 \) and \( \Phi_2 \).

6. Find the best matched domain block from the search spaces \( \Phi_1 \) and \( \Phi_2 \). Record the fractal code of \( \nu_j \).

4. Experimental Results

The proposed search strategy on the frequency domain for fractal image compression is simulated to verify its performance. The tested image is Lena. The image size is \( 256 \times 256 \) and the size of range block is \( 8 \times 8 \). The simulation programs implemented by using Borland C++ Builder v. 6.0 are run on a Pentium Core2 Duo 2.0GHz, 1G RAM, windows XP PC. The quality measurement of the retrieved image \( g(i,j) \) is Peak Signal to Noise Ratio (PSNR) defined as

\[
\text{PSNR} = 10 \times \log_{10} \left( \frac{255^2 \times m \times n}{\text{MSE}(f,g)} \right)
\]

where \( m \times n \) denotes the image size and \( f(i,j) \) is the original image.

![Fig. 4. PSNR of retrieved image versus \( \Delta \theta \) for phase angle span search method.](image)

Fig. 4 shows the relation of PSNR of retrieved image versus \( \Delta \theta \) for proposed method. The figure reveals that, at the beginning of \( \Delta \theta = 0.1^\circ \), the PSNR is 27.02 dB. When \( \Delta \theta \) increases, the image quality also increases accordingly. However, the amount of MSE computations is also increasing and so is the encoding time. The relations of the number of MSE computations and encoding time versus \( \Delta \theta \) are shown in Fig. 5 and 6, respectively. At \( \Delta \theta = 0.8^\circ \), the number of MSE computations 8,919,816 and the encoding time is 42.77 seconds which is about 48 times faster that the full search method. The image quality is 28.02 dB which is only 0.89 dB decay.

Table 1 lists the comparative results of the full search method and proposed method. As observed, the encoding speeds of phase span search method is 58.53 times faster than that of the full search method while there is only 1 dB decay in retrieved image quality.

![Fig. 5. The number of MSE computations versus \( \Delta \theta \) for phase angle span search method.](image)

Table 1. The comparison of full search method and proposed method.

<table>
<thead>
<tr>
<th>Method</th>
<th>( \Delta \theta )</th>
<th>Time (sec)</th>
<th>MSE Computations</th>
<th>PSNR (dB)</th>
<th>Speedup ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full Search</td>
<td>( \times )</td>
<td>2047.30</td>
<td>475,799,552</td>
<td>28.91</td>
<td>1.00</td>
</tr>
<tr>
<td>proposed</td>
<td>0.68°</td>
<td>34.98</td>
<td>7,592,768</td>
<td>27.92</td>
<td>58.53</td>
</tr>
</tbody>
</table>
5. Conclusion

In this paper, we construct a coordinate system using two DCT coefficients of image blocks: the lowest vertical coefficient $F(1,0)$ and the lowest horizontal coefficient $F(0,1)$. We thus map all of the image blocks into the coordinate system with the property that blocks having similar edge shapes will concentrate in some specific regions. A phase span search method is presented to execute the optimal search at the coordinate system. In the encoding process, the searches are limited in the confined regions to find best matched. The experiment demonstrates that the encoding speed of the proposed method is $91.7 \sim 184.5$ times faster than that of the full search method with the cost of only $0.4 \sim 1.0$ dB decay for image quality.

References


