Fractal Image Compression Using Iterated Function System With Probabilities *

Suman K. Mitra, C. A. Murthy, Malay K. Kundu, and Bhargab B. Bhattacharya
Indian Statistical Institute
203, B. T. Road, Calcutta 700035, India
{skmitra_t, murthy, malay, bhargab}@isical.ac.in

Tinku Acharya †
Digital Imaging and Video Division
Intel Corporation, Chandler, AZ 85226, USA
tinku.achrya@intel.com

Abstract

This article deals with a new technique of fractal image compression based on the theory of iterated function system (IFS) with probabilities. The theory of IFS with probabilities, in the context of image compression is a relatively unexplored area. The rationale behind using this approach stems from the fact that it is possible to define a Markov operator associated with the probability measure whose support is the support of the given image. A new technique of fractal image compression is proposed using IFS with probabilities. The technique is found to be extremely fast in computing the coefficients of maps and the probabilities. Thus, the proposed technique provides a very fast fractal-based image compression encoding.

Index Terms : *Contractive map, iterated function system (IFS), Collage theorem, measure, Markov operator, image compression.*

1 Introduction

A set of contractive maps on a space is called iterated function system (IFS) on the same space. The most interesting property of IFS is that it produces a fixed point when the maps are used recursively starting with any arbitrary set. The fixed point is called the attractor of the IFS. The process of finding out contractive maps is thus the main area of interest. In most of the published research papers, implementation part of the theory of IFS with probabilities to real life images is not very clear as the demonstrations are mostly restricted to binary images. Seemingly, a photocopying machine has been designed by means of which coefficients of maps are computed [2]. The concept of photocopying machine has also been extended to the case of grayscale images [2]. In this case, the machine is computing not only the coefficients of maps but also the probability values associated with each map. The IFS with probabilities defines a Markov operator associated with the probability measure whose support is the support of the given image. It has also been shown that the Markov operator is a contractive map on the space of all probability measures. The Collage theorem for measures has also been proposed by Barnsley [2]. From this Collage theorem it has been found that an IFS with probabilities can produce an invariant measure such that the support of the invariant measure is the unique fixed point of the IFS under consideration.

The proposed algorithm is not exactly in the line of the conventional Partitioned Iterative Function System (PIFS) based fractal image compression technique[6, 7] but as mentioned is based on finding out the associated probabilities corresponding to each contractive map. So far, there are only a few attempts have been made to explore the idea of iterated function system with probabilities for image compression. Actually, we have come across very few reference [3, 4, 5] in the literature on fractal image compression, besides the work done by Barnsley [2] on this particular aspect.

In our proposed algorithm we have used a multiscaling division of the given image up to a predetermined level or up to that level at which no further division is required. At each level the coefficients of maps and the corresponding probabilities are computed using the gray value information contained in the image level higher to the level under consideration. Note that the computational procedure of the present algorithm is different than that of Dudbridge [3] and

*This work is funded by a grant from Intel Corp., USA (PO # CAC042717000)

†Author for correspondence

0-7695-1062-0/01 $10.00

others. The algorithm appears to be very efficient in terms of computational time (order of seconds). On the other hand algorithms suggested by Jacquin [6] and Mitra et al. [7] usually take much higher computational time (order of hours) to compute the same for a 256 × 256, 8bit/pixel image. The main objective, in this article, is to investigate the art of finding probabilistic Markov operators for approximating the given image. The high speed of execution of the algorithm appeared to be the main advantage of the proposed algorithm.

The mathematical foundation of IFS with probabilities is outlined in Section 2. The methodology of the proposed algorithm is described in Section 3. Section 4 presents implementation and the results. Discussion and conclusions are provided in Section 5.

2 Mathematical Foundation of IFS with Probabilities

The detailed mathematical description is given in [1, 2]. Some relevant definitions and theorems are stated here. Proofs of the theorems have not been stated as all these are already given in [1]. Starting with the definition of complete metric space, IFS has been defined in the space of probability measures. Invariant measure of IFS in terms of Collage theorem has also been described in the same space.

The most important result which is being used here, is the theorem of invariant measure of IFS with probabilities defined on a probability space. Other definitions and theorems are stated for better understanding of the theorem of interest. Let us start with complete metric space.

Definition 1: *Let (X, d) be a complete metric space. Then $\mathcal{H}(X)$ denotes the space of all nonempty compact subsets of X.* ♠

Definition 2: *The map $w : X \to X$ is called a contractive map with contractivity factor s $(0 \leq s < 1)$ if*

$$d(w(x), w(y)) \leq s\, d(x, y) \;\; \forall x, y \in X. \spadesuit$$

Definition 3: *An iterated function system (IFS) consists of a complete metric space (X, d) together with a finite set of contractive maps $w_i : X \to X$, with respective contractivity factors s_i, for $i = 1, 2, \cdots, n$. The notation for this IFS is $\{X; w_1, w_2, \cdots, w_n\}$ and its contractivity factor is s where $s = max\{s_1, s_2, \cdots, s_n\}$.* ♠

Next we are going to state the Collage theorem of IFS.

Theorem 2 (Collage theorem) : *Let (X, d) be a complete metric space. Let $T \in \mathcal{H}(X)$ and let $\epsilon \geq 0$ be given. Choose an IFS $\{X; w_1, w_2, \cdots, w_n\}$ with contractivity factor s such that*

$$h(T, \bigcup_{i=1}^{n} w_i(T)) \leq \epsilon,$$

where h is the Hausdorff metric. Then

$$h(T, A) \leq \frac{\epsilon}{1-s},$$

where A is the attractor of IFS. ♠

Now we shall define IFS defined on the space of all probability measures.

Definition 4: *An iterated function system with probabilities consists of an IFS $\{X; w_1, w_2, \cdots, w_N\}$ together with an ordered set of numbers $\{p_1, p_2, \cdots, p_N\}$, such that*

$$p_1 + p_2 + \cdots + p_N = 1 \;\; and \;\; p_i > 0 \;\; \forall i.$$

The probability p_i is associated with w_i. ♠.

Theorem 3: *Let $\mathcal{P}$ denote the set of all probability measures on $\square$ and let d_H denote the Hutchinson metric. Then $(\mathcal{P}, d_H)$ is a complete metric space.* ♠

Definition 5: *Let $\{\square; w_1, w_2, \cdots, w_2; p_1, p_2, \cdots, p_n\}$ be an IFS with probabilities. The Markov operator associated with the IFS is the function $M : \mathcal{P} \to \mathcal{P}$ defined by*

$$M(\nu) = p_1 \nu \circ w_1^{-1} + p_2 \nu \circ w_2^{-1} + \cdots + p_n \nu \circ w_n^{-1};$$

$$\forall \nu \in \mathcal{P}, \;\; i.e. \;\; M(\nu) = \sum_{i=1}^{n} p_i \nu \circ w_i^{-1}. \spadesuit$$

The preceding definitions and results are needed to prove the following theorem.

Theorem 4 (Hutchinson's Theorem) : *Let $M : \mathcal{P} \to \mathcal{P}$ be the Markov operator associated with an IFS with probabilities, where each transformation has contractivity factor $0 \leq s < 1$. Then M is a contractive map, with contractivity factor s, with respect to the Hutchinson metric $\mathcal{P}$; i.e.,*

$$d_H(M(\nu), M(\mu)) \leq s\, d_H(\nu, \mu); \;\; \forall \nu, \mu \in \mathcal{P}.$$

In particular, there is a unique measure $\mu \in \mathcal{P}$ such that $M\mu = \mu$. Also, if $M^N(\nu) = M(M^{N-1}(\nu))$ for $\nu \in \mathcal{P}, N \geq 2$, then

$$\lim_{N \to \infty} M^N(\nu) = \mu, \;\; \forall \nu \in \mathcal{P}.$$

where the convergence is with respect to the Hutchinson metric on $\mathcal{P}$. ♠

The Collage theorem for measures is stated below. The theorem states that any given probability measure is approximated by the fixed point of an IFS with probabilities.

Theorem 5 (Collage Theorem for Measure) : *Let $\{\square; w_1, w_2, \cdots, w_n; p_1, p_2, \cdots, p_n\}$ be an IFS with probabilities. Let μ be the associated invariant measure. Let $s \in (0, 1)$ be the contractivity factor for the IFS. Let $M : \mathcal{P} \to \mathcal{P}$ be the associated Markov operator. Let $\nu \in \mathcal{P}$. Then*

$$d_H(\nu, \mu) \leq \frac{d_H(\nu, M(\nu))}{(1-s)}. \spadesuit$$

Note that the theory of IFS with probabilities can be extended to include a *condensation measure*. A condensation measure is a Borel measure μ_0 supported on $\square$ with

$$|\mu_0| = \int_{\square} d\mu_0(x) \leq 1.$$

An IFS with probabilities and condensation measure takes the form

$$\{\square; \mu_0, w_1, w_2, \cdots, w_n; p_1, p_2, \cdots, p_n\},$$

with

$$|\mu_0| + p_1 + p_2 + \cdots + p_n = 1 \text{ and } p_i > 0 \ \forall i.$$

The associated Markov operator is now defined by

$$M(\nu) = \mu_0 + p_1 \nu \circ w_1^{-1} + p_2 \nu \circ w_2^{-1} + \cdots + p_N \nu \circ w_N^{-1}; \quad \forall \nu \in \mathcal{P}.$$

Here μ is the invariant measure of the IFS with probabilities

$$\{\square; \mu_0, w_1, w_2, \cdots, w_n; p_1, p_2, \cdots, p_n\}.$$

Here

$$M(\mu) = \mu_0 + \sum_{i=1}^{n} p_i \mu \circ w_i^{-1} = \mu$$

$$\text{and } \lim_{N \to \infty} M^N(\gamma) = \mu, \ \forall \gamma \in \mathcal{P}.$$

$$\text{where, } M^N(\gamma) = M(M^{N-1}(\gamma)), \quad \forall \gamma \in \mathcal{P} \text{ and } \forall N \geq 2.$$

Thus, once we have the IFS with probabilities or the Markov operator associated with ν in our hand, we can approximate the measure ν by the invariant measure μ starting from any measure γ on the same space.

In the next section the methodology of the proposed scheme is discussed.

3 Multiscale Probabilistic Approach for Image Encoding

In the present algorithm we have used a multiscaling division of the given image. At each level of partitioning, the image or subimage is divided into four quadrants. Each quadrant is called the child of its parent image or subimage. The Markov operator is computed and the approximation of the image by means of the invariant measure of the Markov operator. The process of partitioning the image, or the subimage as the case is, and computing the Markov operator are performed up to a predetermined level or up to a level at which no further division is required. No further division is required at any level indicates the successful approximation of the given image. To compute the Markov operator, it is essential to find the maps and their corresponding probabilities. At each level, probabilities are computed from the proportion of relative image information based on some image features, computed from gray level values, contained in a child subimage and its parent. Actually, contractive maps along with the probabilities try to approximate the proportion of image information contained in a region relative to the total image information. So, it is essential to store the total image information. Also the corresponding contractive maps are nothing but the maps from a parent subimage to its children subimages. The size of a parent subimage is double that of its children subimages. Hence, the map from parent subimage to its child is a contractive map.

Now if the complexity of the pixel arrangement of the parent subimage is higher, then further division of the subimage is likely to be necessary. On the other hand if the image region where the complexity of the pixel arrangement is low, further division may or may not be required. More is the division of the image blocks less will be the efficiency of the encoding process in terms of compression ratio. In order to overcome this problem we have introduced a simple classification scheme. The scheme classifies the children subimages into two groups according to the variability of the pixel values within a block (subimage). If the variability is low *i.e.*, if the variance of the pixel values in the subimage is below a fixed value, called threshold, we call the subimage as smooth type. Otherwise we call it a rough type. Each pixel value in a smooth type subimage is replaced by the mean of all pixel values. The procedure can be looked upon as a condensation mapping. Rough type subimages are approximated by the computed Markov operators. Hence, for each subimage we are to store either a probability value or the mean of the pixel values along with the information regarding the block type.

The complete description of an image in terms of its code provides the compression ratio or bit rate. The important factors which are to be considered during the proposed encoding of an image are : (i) the description of the image partition, (ii) the nature of the blocks, *i.e.*, class information of blocks and (iii) the quantized values of the numerical parameters.

First of all, the sum of the gray level values or the average gray level value of the given image is to be stored. If the given image is a 8 bit/pixel image then 8 bits are required to store this average gray level value. Next a complete description of the image partition is stored. Note that in a multiscaling division, a block is either partitioned into its children blocks or remains unchanged. The partitioning of a block into its children blocks depends on the decision of the estimator which checks the parent block. In particular, for a parent block, there may exist one or two or three or all four children blocks. So, there will be sixteen possible coding configurations for a block at any stage and 4 bits are required for storing this information. again, for storing the characteristics of blocks (either smooth or rough), we need to store another bit for each block. Now, we are left with two remaining most important numerical parameters, the mean gray level value of a smooth type block and probability corresponding to each block. The mean values and the probabilities are quantized and stored in the codes.

The image reconstruction is done by decoding of the code which is obtained during encoding. Here the decoding process is almost similar to the encoding process. The process starts with an arbitrary image. This starting image is playing a role of a dummy image. In the decoding process no probabilities or contractive maps are computed. The starting image is partitioned following the partitioning configuration stored in the codes. Using the block information, the mean gray level values or the stored probabilities and the maps are applied to the blocks of the starting image to get the fixed point for each pixel of the starting image. These fixed points are not the same fixed point of the target image. The desired image *i.e.*, the close approximation of the target image is obtained by multiplying each fixed point of the starting image by the ratio of sum of the gray level values of the original image, stored in the form of codes and that of the starting image.

Note that, the number of iterations to get fixed point of a pixel value of the starting image is fixed, depending on the image size. Also, unlike the decoding of a PIFS code [6, 7], no intermediate reconstruction of the image is possible in this decoding scheme.

4 Implementation and Results

The algorithm has been tested with several grayscale images. We are presenting here the results obtained for a 256 $\times$ 256, 8 bits/pixel "Lena" image. The image is subdivided into four, 128 $\times$ 128 subimages, each of which is encoded separately. During encoding, the partitioning of each subimage has been carried out up to a maximum level where subimages are of size 2 $\times$ 2.

For comparison, the results of both partitioning up to 2 $\times$ 2 subimages and 4 $\times$ 4 subimages are given here. The encoding with partitioning up to 2 $\times$ 2 subimages for different threshold values are also studied. All test results and some statistics are given in Table 1. The test result shows how compression ration increases with decrease in the PSNR value for encoding up to 2 $\times$ 2 subimages. The time taken for encoding is also computed to show the fastness of the proposed compression scheme. All the programs are tested on a 133 MHz Silicon graphics (Indy) Workstation.

The results for the "Lena" image (Figure 1) are shown in figures 2 to 4. Figures 2, 3 and 4 are showing the decoded "Lena"for different threshold values when the encoding is processed up to the subimages of size 2 $\times$ 2. The decoding process is carried out with a starting "White" image with all pixels having gray level value 255.

Table 1. Test results for 256 $\times$ 256, 8 bit/pixel "Lena" image

Lowest level subimage size	Compr-ession ratio	PSNR (in db)	bits/ pixel (bpp)	Time elapsed (in Sec.)
4 $\times$ 4	20.1	22.43	0.40	8.34
2 $\times$ 2	3.20	30.49	2.50	23.39
2 $\times$ 2	5.08	28.23	1.58	16.40
2 $\times$ 2	8.81	27.11	0.90	10.93

The proposed multiscale probabilistic approach for fractal image compression is compared, in terms of the computational time, with the GA based fractal image compression technique [7]. The basic feature of GA based fractal image compression scheme is same as that of the scheme proposed by Jacquin [6] and is usually known as partitioned iterated function system (PIFS) based compression. The basic aim of the GA based compression technique is to reduce the computational time for the implementation of the PIFS based compression technique. It is found that at least 20 times reduction in the search space is achieved in the GA based compression technique [7] in comparison to the usual PIFS based image compression technique using exhaustive search. For comparison, the test results of both probabilistic approach encoding and the GA based encoding techniques are presented in Table 2.

Table 2. Results obtained by using multiscale probabilistic approach technique and the GA based technique for fractal image compression

Image	Compression Ratio	PSNR (in db)	Time elapsed (in Sec.)
Lena	probabilistic approach		
	8.81	27.11	10.93
	GA based technique		
	10.50	30.22	3141.47

It is very clear from the Table 2 that the proposed multiscale probabilistic approach for fractal image compression is very fast compare to the GA based fractal image compression technique. It is true that compression ratio and the quality of the decoded image in terms of PSNR of the proposed method is not so impressive but they are comparable with those of other fractal based image compression techniques [6, 7, 8, 9]. Blocking effects in some small number of regions are visible in the decoded image due to quantization. But in terms of computational time the present method is probably one of the fastest methods proposed so far.

5 Conclusion

The proposed algorithm runs very fast, but its performance strongly depends on the partitioning rule used. The present technique based on probabilistic approach for fractal image compression is at least 300 times faster than the GA based fractal image compression technique [7]. Hence, it may be convenient for real-time video coding. This work is currently under investigation.

Figure 1. Original 256 × 256, 8 bpp Lena image

Figure 2. Decoded Lena, 2.50 bpp

Figure 3. Decoded Lena, 1.58 bpp

Figure 4. Decoded Lena, 0.90 bpp

References

[1] M. F. Barnsley. *Fractals Everywhere.* Academic Press, New York, 1988.

[2] M. F. Barnsley and L. P. Hurd. *Fractal Image Compression.* AK Peters Ltd., Massachusetts, 1993.

[3] F. Dudbridge. Linear time fractal coding schemes. In Y. Fisher, editor, *Fractal Image Encoding and Analysis, NATO ASI Series F, vol. 159*, pages 133–137. Springer Verlag, Berlin, 1998.

[4] B. Forte and E. R. Vrscay. Solving the inverse problem for measures using iterated function systems : a new approach. *Adv. appl. Prob.*, 27:800–820, 1995.

[5] S. Graf. Barnsley's scheme for the fractal encoding of images. *Journal of Complexity*, 8:72–78, 1992.

[6] A. E. Jacquin. Image coding based on a fractal theory of iterated contractive image transformations. *IEEE Transactions on Image Processing*, 1(1):18–30, 1992.

[7] S. K. Mitra, C. A. Murthy, and M. K. Kundu. Technique for fractal image compression using genetic algorithm. *IEEE Transactions on Image Processing*, 7(4):586–593, 1998.

[8] L. Thomas and F. Deravi. Region-based fractal image compression using heuristic search. *IEEE Transactions on Image Processing*, 4(6):832–838, 1995.

[9] C. J. Wein and I. F. Blake. On the performance of fractal compression with clustering. *IEEE Transactions on Image Processing*, 5(3):522–526, 1996.