

# OPTIMIZED BLOCK SIZE BASED VIDEO CODING ALGORITHM

Gayatri<sup>1</sup>, Shwetha.V<sup>2</sup>

<sup>1</sup>Post Graduate, Electronics and Communication Department, SJCIT Chikballapur, Karnataka, India

<sup>2</sup>Assistant Professor, Electronics and Communication Department, SJCIT Chikballapur, Karnataka, India

## Abstract

Multi-view video coding is a new scope in the technology that is of the utmost important in the stage where 3D TV, Free View Video(FVV), Free View point Television, Video Conference are in its peak research.MPEG/VCEG jointly have come up with the compression of video sequence taken from multiple cameras during the single video stream. This compression would largely help the communication channels, which deal with the video transmission and delivery .It must adapt compression techniques as it contain large amount of data which has to transmitted and stored .This project deals with new compression technique called the block based fractal compression technique. Fractal compression technique has been modified to block sized fractal compression in order to bring more accuracy measures and also higher compression ratio. The block based term here means that the block which are classified as priority blocks or non-priority. Priority of the region is considered on basis of information in the pre-processing. The main aim is to reduce redundancy problem by considering information of regions. Matlab based simulation is carried out and the results of this method is measured and tabulated.

**Keywords:** H.264/AVC, multi-view video coding, fractal, redundancy, macro block, disparity.

\*\*\*

## 1. INTRODUCTION

Video coding is a process of compressing and decompressing a digital video signal. Visual communication is one of the important evolving technology that enhance everybody's interest in fields such as 3D image processing, Holography and Multi View Video MVV. Multi-view coding systems are being proposed for new multimedia service like Free Viewpoint Television (FVT), Viewpoint Video (FVV), Video Conferencing, 3DTV.

A set of video sequences captured from different view angles of the camera of a particular same scene [1-2], among all views there exists a large amount of correlation but due to this feature there exists demand for extensive storage and transmission bandwidth due to the massive data volume. This scheme of exploit of extra redundancies among inter-view and temporal statistical dependencies incurs extensive computation. So we must concern on information that multi-view video coding must exploit the redundancies among the views. But extraction of extra redundancies lead to large computations which benefit by giving gained coding efficiency. The other benefits regarding this are decreased computational and memory complexity, efficient random access in time for certain views, delay management ,buffer management, as well as parallel processing. Temporal prediction is an important mechanism to remove redundancy .Correlations between adjacent cameras are exploited in addition to temporal correlations with in each sequence . Therefore multi-view coding adds an another compression dimension on top of single view coding i.e, inter-view direction[2].Disparity and motion estimation leads to

computational complexity ,therefore its good to low the computation complexity while emphasizing the relation between them[3,4].The best quality of MVC is that it have high degree of temporal statistical correlations between temporally succeeding images. Prediction structure using hierarchical B pictures is adopted by Joint Video Team (JVT). H.264/AVC follow block based coding techniques where temporal and view correlation with in temporally successive pictures and neighboring views[5].

Low coding efficiency is absorbed by adopting schemes which uses spatial prediction or temporal prediction mode as they not sufficiently exploit the correlation between views. So only fast disparity and motion estimation method is adopted[6].By adopting different methods prediction vectors are obtained, by there difference reliability of each macro block is calculated [7,8].Realization of MVV encoder is very difficult in real time applications as there is huge & extensive amount of video data and due to this it leads to high computational complexity.

We adopt H.264/AVC codec for the video compression, where compression is based on the fractal algorithm. Here there is meaningful partition of blocks as intra and inter blocks based on important part of input data. One of the most important drivers for the standardization of H.264 and its subsequent adoption by industry is its improved performance compare to earlier standards. The prediction method of H.264/AVC are more flexible and leads to reduction in complexity of motion estimation and disparity.

As fractals are self similar or recursively transformed copies of either themselves or part of themselves. These poses high

degree of redundancy and as H.264/AVC has better prediction quality which leads to be more compatible with fractal concept leading to better results.

The rest of paper organized as follows: Section II explains about fractal coding theory. Section III explains about adopted partition scheme and Search strategies. Section IV explains reason for H.264/AVC adoption. Section V explains about experimental results and section VI conclusion of this paper. VII Explains about future enhancement.

## 1.1 Fractal Coding Theory

Fractal coding method is a novel compression method, in which the self-similarity of the video is used to eliminate redundancy existed in video data. Due to its innovative idea, high compression ratio, high decoding speed, independent resolution etc it receives wide research attention. The idea of fractal image coding[11-12] is based on Iteration Function System in which the governing theorem is the Contractive Mapping Fixed Point theorem.

**Contractive Maps:** Let  $X$  be a metric space with metric  $d$ . A map  $\omega: X \rightarrow X$  is with Lipschitz factor  $S$  if there exists a positive real value  $s$  such that

$$d(\omega(x), \omega(y)) \leq s d(x, y)$$

**The Contractive Mapping Fixed-Point Theorem:** Let metric space and  $f: X \rightarrow X$  be a contractive mapping.

Then there exists a unique point  $x_f \in X$  such that for any point  $x \in X$ .

$$x_f = f(x_f) = \lim_{n \rightarrow \infty} f^{on}(x)$$

such a point is called fixed point or the attractor of mapping  $f$ .

**Generalized Contraction Theorem:** Let  $f$  be eventually contractive with exponent  $n$ . Then there exists a unique fixed point  $x_f \in X$  such that for any  $x \in X$ .

$$x_f = f(x_f) = \lim_{n \rightarrow \infty} f^{on}(x)$$

In that case,

$$d(x, x_f) \leq \frac{1}{1-s} \frac{1-\delta^n}{1-\delta} d(x, f(x))$$

Where  $S$  is the contractility factor of  $f^{on}$  and  $\delta$  is Lipschitz factor  $f$ .

Let  $X$  be a complete metric space. An iterated function system is a collection of contractive map  $\omega_i$ :

$$X \rightarrow X \text{ for } i=1, \dots, n.$$

Fractal algorithms convert these parts into mathematical data called "fractal codes" which are used to recreate the encoded image.

A metric space is an ordered pair  $(M, d)$  where  $M$  is a set and  $d$  is a metric on  $M$  i.e, is a function.

$$d: M \times M \rightarrow \mathbb{R}$$

Such that for any  $x, y, z \in M$ , the following holds:

1.  $d(x, y) \geq 0$  (*non-negative*)
2.  $d(x, y) = 0$  if  $x = y$  (identity of indiscernible)
3.  $d(x, y) = d(y, x)$  (Symmetry)
4.  $d(x, z) \leq d(x, y) + d(y, z)$  (Triangle inequality)

The first condition follows from the other three, since for any  $x, y \in M$ .

1.  $d(x, y) + d(y, x) \geq d(x, x)$  (by triangle equality)
2.  $d(x, y) + d(x, y) \geq d(x, x)$  (By symmetry)
3.  $d(x, y) \geq 0$  (By identity of indiscernible)

The function  $d$  is also distance function or simply distance often  $d$  is omitted and one just writes  $M$  for a metric space if it is clear from the context what metric is used.

## 1.2 Classification of Fractal Coding based on following Principles

1. Partition imposed on the image support by the range blocks.
2. Composition of pool of domain blocks.
3. Class of transforms applied to the domain blocks.
4. Type of search used in locating suitable domain blocks.
5. Representation and quantization of the transform parameters.

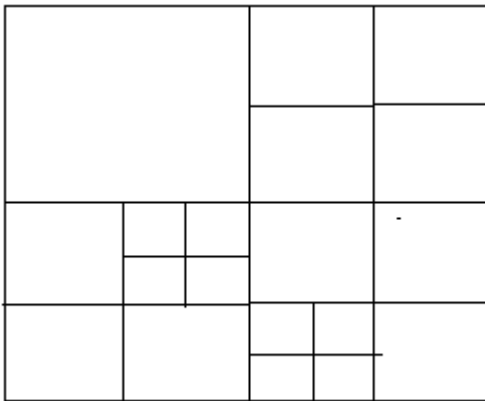
## 2. PROPOSED METHOD FOR PARTITIONS SCHEMES AND SEARCH STRATEGIES

Choice in type of image partition for range blocks is considered as vital and is considered as initialization step for the designing a fractal coding scheme. There is restriction to the possible sizes and shape of domain blocks as these blocks must be transformed to cover range blocks and the type of transformation too will effect in considering sizes and shape of domain blocks. Majority of blocks considered are rectangular in shape.

## 2.1 Quadtree

Fig-1 Quadtree is a tree data structure in which each internal node has exactly four sub nodes. Quadtree are most often used to partition two dimensional space by recursively subdividing it into four quadrants or regions. The regions may be square or rectangular, or may have arbitrary shapes. Features of Quadtree are :-

- They decompose space into cells.
- Each cell has a maximum capacity, when maximum capacity is reached, the cell splits.
- The tree dictionary follows the spatial decomposition of the Quadtree.



**Fig-1 Quadtree decomposition**

The Quadtree decomposition builds a tree. Each tree node has four sub nodes and it is associated with a uniquely defined region of the image. It is obvious that the root is associated with the whole image. Quadtree decomposition can be done either by top-down or bottom-up procedures.

Fractal compression algorithms require lengthy coding times for computational requirements of the domain search. Efficient search of domain search techniques require much more research in fractal coding.

## 2.2 Search Strategies

The transformed version of domain block should match the range block and for this procedure it have to search for the best domain block for particular range block. But the problem is to find for each range block, the domain block that can be made the closest by admissible transform.

To simplify this problem each image block is invariantly represented. Here both range blocks and domain blocks are transformed and by distance comparisons of these blocks the best possible match is determined. Here we apply DCT transformation on both range and domain blocks followed by normalization. Here search is carried out at each level as to find the best match .Let us assume set of domain blocks  $d_i$  and there transform as  $M_p$  parameterized by  $p$  the ,the domain

block for particular range block results in collage error of minimum  $p$  is equal to  $\|r - M_p d_i\|$ .

In DCT fractal algorithm following algorithm is followed,

## Encoding

1. In the quad tree approach, the fractal dct encoding is done in four levels.
2. In level 1, define the error for distance metric and Normalization matrix for quantizing dct output for better compression.
3. In level 1,take range blocks size  $16*16$  and domain blocks of size  $32*32$
4. Read the input gray scale image.
5. Take all possible range blocks and apply DCT transform and divide with the normalized matrix and save all the transformed range blocks to Some variable.
6. Take all the possible domain blocks with overlapping and calculate average matrices of size  $16*16$  by taking average of neighbor pixels. Apply DCT transform for average matrix and Divide with the normalized matrix and save all the transformed domain blocks to some variable.
7. Take each transformed range block from TRR and match with the all the transformed domain blocks from TDD by using distance metric
8. Distance metric is obtained by taking difference of those matrices and multiply with defined contrast matrix and obtain the distance and save all the distances
9. Take the domain block for which distance is less than defined error and replace it in the Range block position and save the best contrast and distance.
10. In level 2 take range blocks which are not getting best matched domain block in the distance metric and redefine size of range block and domain block to half of the size in the level1.
11. In level 2 redefine the error and normalized matrix and repeat the procedure as that in level 1 for new range and domain block sizes .
12. Repeat the procedure in level2 for level3 and level 4 for those range blocks whichever is not getting best matched domain block.

## Decoding

Decompression is done by applying the transformations on an arbitrary image iteratively until an image produced looks approximately like the original image.

## 2.3 Macroblock Partitions

Complexity in calculations arises when a frame is partitioned by quadtree partition and the iteration is used in matching process. Calculation speed and complexity of video

compression algorithm is largely affected in positive way by the adoption of macroblock partition. H.264's macroblock partition is used in this paper.

The best match for range block is determined by evaluating RMS (Root Mean Square). The minimum RMS means the matching. It is determined by

$$\text{RMS} = \frac{1}{N} [\sum_{i=1}^N r_i^2 + s(\sum_{i=1}^N d_i^2 - 2 \sum_{i=1}^N r_i d_i + 2o \sum_{i=1}^N d_i^2) + o(N \cdot 2 \sum_{i=1}^N r_i)]$$

Where the contrast factor S is

$$s = \frac{\sum_{i=1}^N r_i d_i - \sum_{i=1}^N r_i \sum_{i=1}^N d_i}{\sum_{i=1}^N d_i^2 - (\sum_{i=1}^N d_i)^2}$$

and the offset factor o is

$$o = \frac{1}{N} [\sum_{i=1}^N r_i - s \sum_{i=1}^N d_i]$$

$r_i$  is pixel value of Range block (R),  $d_i$  is pixel value of Domain block (D),  $N$  is the number of pixel in block. Before the block matching processing, RMS of the whole macroblock is calculated.

### 3. H.264/AVC

It is an industry standard; it defines a format for compressed video data; it provides a set of tools that can be used in a variety of ways to compress and communicate visual information; it is a stage in an involving series of standardized methods for video compression.

Previously popular standards used 8x8 discrete cosine transform but in H.264 we use 4x4 transforms that can be computed exactly in integer arithmetic. Because of advent the integer arithmetic it avoids inverse transform mismatch problems. Here in this new transforms, multiplication is not involved instead just additions and shifts, in 16-bit arithmetic, thus minimizing computational complexity, especially for low-end processors. By using short tables, the new quantization formulas use multiplications but avoid divisions. The new features which are added to H.264 are advance intra coding, enhanced motion estimation with variable block size, multiple reference pictures, integer block transform, improved deblocking filter, entropy coding, and error resistant coding.

An H.264 video encoder carries out prediction, transforming and encoding processes to produce a compressed H.264 bit stream. An H.264 video decoder carries out the complementary processes of decoding, inverse transform and reconstruction to produce a decoded video sequence.

Here intra (prediction in current frame) and inter prediction is carried out prediction or from other frames that have already been coded and transmitted using inter prediction. H.264 employs a 4x4 integer transform:

$$H = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & -2 \\ 1 & -1 & -1 & 1 \\ 1 & -2 & 2 & -1 \end{bmatrix} \quad \text{---}$$

Here the transform is an approximation of the DCT, It has a similar coding-gain to the DCT transform, Since the integer transform has an exact inverse operation, there is no mismatch between the encoder and the decoder which was a problem in all DCT based codec's.

Before transmission, the generated data of all types are entropy coded. H.264 supports two different methods of the entropy coding

1. Context adaptive variable length coding (CAVLC)
2. Context adaptive binary arithmetic coding (CABAC)

### 4. EXPERIMENTAL RESULTS

The compression ability of the fractal based methods for images and video is highly efficient. In multi-view video containing background such as tree, road, mountain, houses and so on. Here according to the information of the regions and setting priority among regions we can solve the redundancy problems. Here priority of the regions is decided by information in the pre-processing.



Fig-2: Input Video snapshot



Fig-3: Encoded Video snapshot



Fig-4: Decoded Video snapshot

#### 4.1 Evaluation Result

Objective measure widely used is peak signal-to-noise ratio (PSNR) of the signal which is given as

$$PSNR=10 \log_{10} [65025 / MSE]$$

MSE is Mean Squared Error. MSE is found between the original image and decoded video samples. Compression ability of different algorithms are plotted over bit rate of PSNR.

#### 4.2 Expected Results

Two coding tools are simulated i.e, fractal based on variable block-size and full search algorithm to improve efficiency of multi-view video coding . Here Matlab software,7.13 version was used. Here we took 10 frames of a particular video called vodocoder. Here in experiment we consider 10 frames of size 720p x 1080p and video considered is vocoder video.

By comparing the performances of the proposed algorithm and full search algorithm by Table-1 considering input video in terms of processing time and PSNR by this we evaluated the coding complexity. The processing time of proposed algorithm is less compared to full search algorithm. Estimation time is reduced and due to this there is speed up in estimation process.

Table-1: Comparison of processing time

Methods	Test data (vocoder)
Full search algorithm	0.80msec
Proposed method	0.59.65msec
Reduced time(%)	25.4

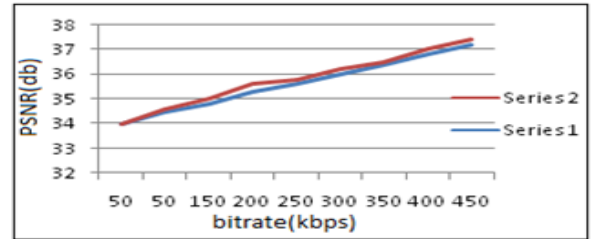


Fig-5 distortion performance of difference multi-view video coding

Series 2= FSA      Series 1= proposed

PSNR gain decrease by 0.1-0.2 dB but compared to the proposed method and full search algorithm there is no subjective change from subjective aspect.

#### 5. CONCLUSIONS

By considering the information of the regions and setting up priority among region based on information we solve the redundancy reduction problem by proposed fast disparity and motion estimation. The computational complexity of search range is reduced by motion estimation and disparity estimation to reduce the search region and the number of reference frame . With same compression quality the encoding time is reduced up to 25.4%.

#### FUTURE ENHANCEMENTS

Homogeneous motion is absorbed in the spatial areas larger than 16x16 for higher resolution sequences like 720px1080p. Therefore motion partition sizes is extended more than 16x16. Motion partition is extended to 64x64.

#### REFERENCES

- [1]. A. Smolic. P. Kauff, "Interactive 3-D Video Representation and Coding Technologies," Proceedings of IEEE, vol.93,pp. 98 ,January 2005.
- [2]. A. Smolic, K. Mueller, N. Stefanoski, J.Ostermann, A.Gotchev,G.B.Akar, G.A. Triantafyllidis and A Koz: "Coding algorithms for 3DTV-A Survey" IEEE Transactions on Circuits and Systems for Video Technology,Vol &,pp. 1606-1621,November 2007.
- [3]. M.Siegel, S.Sethuraman, J.McVeigh, A.Jordan. "Compression and interpolation of 3-D stereoscopic and multi-view video." Stereosc.Disp.Virtual Real.Syst.,pp.227-238,1997.
- [4]. M.Tagliasacchi, L.Friggerio, S.Tubaro, "Rate-distortion analysis of motion-compensated interpolation at the decoder in distributed video coding", IEEE Signal Processing Letters, Vol.14, pp.625-628, September 2007.
- [5]. k.Mueller, P.Merkle, H.Schwarz,T.Hinz, A.Smolic, T.Oelbaum, T.Wiegand,"Multi-view video coding based on H.264/AVC using hierarchical B-frames", Proceedings of the Picture Coding Symposium,Beijing,China,2006.

- [6]. Y.Kim, J.Kim, K.Sohn,"Fast disparity and motion estimation for multi-view video coding", IEEE Trans.Consum.Electron. Vol.53,pp 712-719(2007).
- [7]. M.Flieri and B.Girod,"Multiview video compression",IEEE Signal Process.Mag.,pp.66-76.November 2007.
- [8]. P.Merkle, A.Smolic, K.Muller, and T.Wiegand."Efficient prediction structure for multiview video coding," IEEE trans Circuits Syst VideoTechnology.,vol.17,pp.1461-1473, November 2007.
- [9]. Vetro, P.Pandit. H.Kimata, and A.Smolic,"Joint Multiview Video Model (JMVM5.0),"Joint Video Team, Doc.JVT-X207, Geneva Switzerland, July 2007.
- [10]. Y.T.Kim, J.Y.Kim, and K.H.Sohn,"Fast disparity and motion estimation for multi-view video coding",IEEE Trans Consum.Electron, vol.53,May 2007.
- [11]. A.E.Jacquin, "Fractal Image Coding: A Review",Proc.IEEE.Vol.81,pp.1451-1265,October 1993.
- [12]. B.Wohlberg and G.D.Jager,"A Review of the Fractal Image Coding Literature",IEEE Transactions on Image Processing,Vol.8,pp.1716-1729 December 1999.
- [13]. Wiegand, T.Gray J. Sullivan, Gisle Bjontegaard and Ajay Luthra "Overview of the H.264/AVC Video Coding Standard",IEEE Transactions on Circuits and Systems for the video Technology, July 2003.
- [14]. Narasak Boonthep,Werapon Chiracharit,Kosin Chamnongthai "Variable Block-size Based Fractal Coding for Multi-view Video Coding",IEEE 2013.