

Video Coding Technique Using 3D Dualtree Discrete Wavelet Transform with Multi Objective Particle Swarm Optimization Approaches

¹Thamarai Muthusamy and ²Shanmugalakshmi Ramachandran

¹Department of Electronics and Communication Engineering,
Faculty of Engineering, Karpagam College of Engineering, Coimbatore, India

²Department of Computer Science and Engineering, Faculty of Engineering,
Government College of Technology, Coimbatore, India

Received 2012-12-04, Revised 2013-03-12; Accepted 2013-04-12

ABSTRACT

Wavelet-based image compression with Wavelet transform based motion compensated video codec performs better compression to meet the rate and distortion constraints in video transmission with the available bandwidth than the block based techniques. However, it is well known that the 2D DWT does not represent directional features of images efficiently. Lots of efforts have been contributed to multiscale directional representation. In this study Video coding using Dualtree Discrete Wavelet Transform is considered and its expansive nature is reduced by noise shaping algorithm and high compression ratio is achieved by means of the selection of optimal coefficients using MOPSO method. In Multiobjective Optimization (MO) problems, more than one objective functions have to be minimized simultaneously. In the proposed method, Entropy of the sub bands and Mean square error are considered for optimization with the constraints of frame rate. The two different types of MOPSO approaches such as weighted aggregation approach and Vector Evaluated PSO are used to select the optimum subbands. The two techniques outputs are compared with the standard 3D SPIHT coding.

Keywords: Dual Tree Discrete Wavelet Transform, Noise Shaping Algorithm, Multi Objective Particle Swarm Optimization, Multiobjective Optimization, Mean Square Error, Vector Evaluated PSO

1. INTRODUCTION

The storage and transmission of the video signals are not possible without compression, since video information of one second duration takes several megabytes of memory. The available bandwidth is insufficient for multimedia applications. Research works on compression algorithms have been carried out for video coding. At high compression ratios block based coders introduce artifact and ringing effects. While considering the wavelet based motion compensated 2D+t transform, the complex motion estimation and compensation is a tedious process. So we go for 3D-Wavelet based coders. Discrete Wavelet Transform is commonly used Wavelet Transform (DWT), but having

limitations such as aliasing, oscillatory nature of coefficients and lack of directionality as discussed in Thamarai and Shanmugalakshmi (2011). Dual Tree Complex Wavelet (DTCWT) Transform is more suitable for video coding with the kernel functions having the capability of directional property, eliminating motion estimation and compensation process. Yang *et al.* (2007) reported the suitability of dual tree discrete wavelet transform for video coding.

Dual tree discrete wavelet transform is an expansive type transform; It converts M number of samples into N number of coefficients ($N > M$). The number of coefficients N is reduced using noise shaping algorithm. The significant coefficients are selected using MOPSO algorithm. Thamarai and Shanmugalakshmi (2010)

Corresponding Author: Thamarai Muthusamy, Department of Electronics and Communication Engineering,
Faculty of Engineering, Karpagam College of Engineering, Coimbatore, India

reported the DDWT optimal subband selection using PSO with single objective function. Optimizing the problem with single objective often leads to unacceptable results with respect to the other objective. Therefore a perfect multi objective solution that simultaneously optimizes each objective function is almost impossible. One solution for this multi objective problem is to find a set of solutions which satisfies all the objectives in an acceptable level and which does not dominate any other solution. In this study, video compression is formulated as a multi objective problem and the objective factors identified are MSE, Entropy and computation time with the constraints of frame rate.

1.1. Particle Swarm Optimization and MOPSO

PSO is widely accepted and focused by researchers due to its profound intelligence and simple algorithm structure. Currently PSO has been implemented in a wide range of research areas of functional optimization, pattern recognition, neural network training and fuzzy system control and is successful. In PSO, each potential solution is considered as one particle. The system is initialized with a population of random solutions (particles) and searches for optima (global best particle), according to some fitness function, by updating particles over generations; that is, particles “fly” through the N-dimensional problem search space to find the best solution by following the current better-performing particle. When compared with Genetic Algorithm, PSO has very few parameters to adjust and easy to implement. Binary PSO, Hybrid PSO, Adaptive PSO and Dissipative PSO are variants of PSO and used in various image processing applications.

In PSO, we assume that the problem is in a D-dimensional space, which includes many particles; each particle represents a feasible solution of optimization problem. On every iteration, each particle updates itself by the two extreme values, one is individual extreme value p_{id} , which is personal best value for that particle p_{id} , the other is the global best value for that particle (gbest) p_{gd} . Each particle adjusts its flight speed and direction according to current rate, p_{best} and g_{best} using Equation 1 and 2 repeatedly:

$$v_{id}(t+1) = wv_{id}(t) + c_1 \text{rand}_1(.) (p_{id} - x_{id}) + c_2 \text{rand}_2(.) (p_{gd} - x_{id}) \tag{1}$$

$$x_{id}(t+1) = x_{id}(t) + v_{id}(t+1), 1 \leq i \leq N, 1 \leq d \leq D \tag{2}$$

where, N is the number of particles and D is the dimensionality:

- $V_i = (v_{i1}, v_{i2}, \dots, v_{iD})$, $v_{id} \in [-v_{min}, v_{max}]$ is the velocity vector of particle i, which decides the particle's displacement in each iteration
- Similarly, $X_i = (x_{i1}, x_{i2}, \dots, x_{iD})$, $x_{id} \in [-x_{min}, x_{max}]$ is the position vector of particle i which is a potential solution in the solution space
- The quality of the solution is measured by a fitness function; w is the inertia weight which decreases linearly during a run
- c_1, c_2 are both positive constants, called the acceleration factors which are generally set to 2.0
- $\text{rand}_1(.)$ and $\text{rand}_2(.)$ are two independent random numbers distributed uniformly over the range [0, 1]
- p_{gd}, p_{id} are the best solutions discovered so far by the particle in the group and itself respectively
- $v_{id}(t+1)$ -velocity of the particle at time (t+1) and $x_{id}(t+1)$ -position of the particle at (t+1)

Recently PSO has been extended to deal with multiple objective optimization problems (Parsopoulos and Varahatis, 2002). In the past few years many research works have been focused on modifying PSO to handle multiple objective optimization problems known as multi objective particle swarm optimizer MOPSO. The fixed population size MOPSO and variable population size PSO (Dynamic PSO) are used throughout the evolution process to explore the search space to discover the non dominated individuals (particles). Most of the real life problems are multi objective nature.

Here Video compression is viewed as a multiobjective problem. The fixed population size MOPSO is used throughout the evolution process to explore the search space to discover the non dominated individuals (particles).

1.2. Multi Objective PSO (MOPSO)

PSO is particularly suitable for multi objective optimization because of high speed of convergence. Konak *et al.* (2006) proposed GA for multi objective problems because of its high speed of convergence. But when compared to GA, PSO has memory and very few parameters for adjustment. Chin-Wei and Rajeswari (2010) reported the image segmentation process using Multi objective Optimization. Boukhobza *et al.* (2013) reported the usage of multi objective PSO filter bank design. Coello *et al.* (2004) reported about the PSO algorithm's ability to handle multi objective optimization problems.

Parsopoulos and Varahatis (2002) reported the Extension of PSO for multi objective problem. In order to handle multiple objectives, PSO must be modified before being applied to MO problems. In most approaches, the major modifications of the PSO algorithms are the selection process of g_{best} and p_{best} .

Coello *et al.* (2004) developed a grid based gbest selection process and also employed a second population to store the non dominated solutions. From the second population, using Roulette wheel selection, they selected the gbest randomly. The pbest selected according to the Pareto dominance.

1.3. Vector Evaluated PSO

In Vector evaluated particle swarm optimization VEPSO, n number of swarms are used to solve n objectives. In VEPSO algorithm, when one swarm updates the velocities of the particle, the other swarm is used to find the best particle to follow. In vector evaluated Genetic algorithm-non pareto approach, fractions of the next generation, or sub populations are selected from the old generation according to the objectives separately. After shuffling all the sub populations together, cross over and mutation are applied to generate new populations. In the proposed work, modified weighted aggregate approach as well as VEPSO approach are used to find the optimal subbands of the DDWT. The fitness function of the MOPSO combines the MSE, ESUM-Entropy of the various subbands with different weightage values. In VEPSO the particles are divided into three sub swarms and are used to search for their own objective functions. The objective functions used in both approaches are same.

2. MATERIALS AND METHODS

At first a frame is taken from the video and converted to wavelet coefficients by using Dual tree Discrete Wavelet Transform (DDWT). In the wavelet transform, high frequency components are removed while the low frequency components are preserved to compress the image due to high energy concentration in low frequency subbands. Then Multi-objective PSO is applied to select the optimal subbands. Weighted aggregate approach and VEPSO methods are used to search optimal wavelet subband coefficients of DDWT. The selected optimum subband coefficients are encoded using Variable Length coding algorithm. The block diagram of the proposed system is shown in Fig. 1.

2.1. MOPSO Problem Formulation

Multi objective optimizer finds the optimum solution for the given n number of objective functions with constraints. For the video coding problem, the MOPSO system is formulated as shown in Fig. 2. The inputs are the subband coefficients and outputs are the optimum subbands, which provides minimum MSE and Entropy with less computation time.

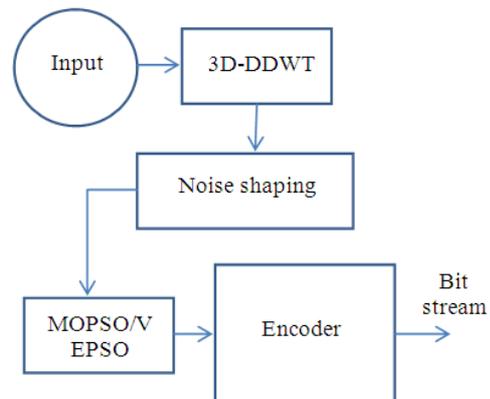


Fig. 1. Block diagram of the coding system

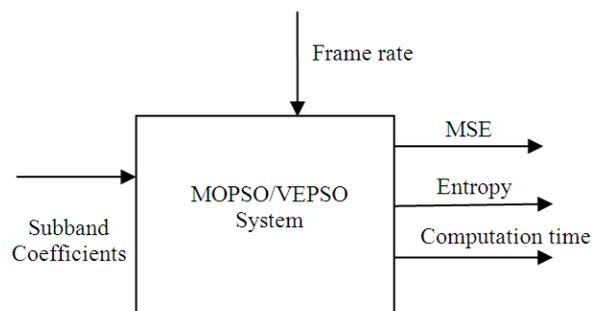


Fig. 2. MOPSO block in the proposed system

Frame rate and bits per pixel are considered as constraints.

2.2. Optimal Subband Selection Using MOPSO Techniques

The two MOPSO approaches are used in this study to select the optimum subbands with objective functions of MSE and Entropy Sum of the Subbands (ESUM).

2.3. Weighted Aggregate Approach

In this MOPSO method, all objective functions are combined as a single function with suitable weightage. The objective functions are minimum Entropy of the subbands, minimum Mean Square Error. The set of optimal subbands are selected using these objective functions.

Ouyang *et al.* (2010) used, MSE and ESUM as objective functions for selecting the best wavelet packet. The MOPSO’s fitness function for video coding is defined as follows Equation 3:

$$\text{Fitness} = a_1\text{MSE} + a_2\text{ESUM} \tag{3}$$

where, a_1 and a_2 are constants and their (weightage values) values are taken as 0.6 and 0.4 respectively. The Entropy sum is calculated for each set of subbands as Entropy of (LL, LH, HL, HH).

$M \times N$ is the size of the image (m,n) , $X^{\wedge}(m,n)$ are the pixel gray value of the original image and reconstructed image respectively.

The computation time varies with respect to the size of the subbands sets and the average Mean Square Error (MSE) σ_{ϵ} is calculated as per Equation (4):

$$\sigma_{\epsilon} = \sum_{l=1}^L \sum_{i=1}^M \sum_{j=1}^N ((I_{org}(i,j,l) - I_{com}(i,j,l))^2) \quad (4)$$

- $I_{org}(i, j, l)$ -Original Image frame in the video sequence
- $I_{com}(i, j, l)$ -reconstructed frame in the video sequence

where, L is the number of frames in the video sequence.

Algorithm:

- Step 1: The video sequence is grouped into Frames. Apply 3D Dualtree wavelet transform and subject the coefficients to noise shaping algorithm. The result in the reduced number of coefficients.
- Step 2: Initialization of population: Set of randomly selected subband coefficients are considered as Particles. Each particle is initialized as the multiplication of randomly initialized 0 or 1 Matrix with the subband of DDWT.
- Step 3: Apply inverse dual tree wavelet transform. The reconstructed image is obtained. Then calculate MSE, Entropy.
- Step 4: Calculate each particle's fitness value according to the Equation No. (3)
- Step 5: If the particle's fitness value is better than the particle's best fitness value, P_{id} is updated. If the fitness value is better than the global best fitness value, P_{gd} is updated.
- Step 6: Similarly, each particle's velocity and position are updated according to the update Equations 1 and 2.
- Step 7: The exploration process continues until a pre specified iterations are satisfied. Declare the global optimum value as the solution.

For the given constraints of frame rate, the PSNR, MSE and the Computation time are measured. According to the weighted aggregate MOPSO approach the best

particle-set of optimum subband coefficient is selected. The average value of ESUM, PSNR and Computation time are also calculated:

- The values taken are $w = 0.9$; $c_1 = c_2 = 2$ $itr = 10$
- The number of particles used: 10, Number of iterations: 10
- PSNR value of the sequence is calculated using Equation 5

$$PSNR = 10 \log \left(\frac{I_{max}^2}{\sigma_{\epsilon}} \right) \quad (5)$$

where, $I_{max}^2 = 255$ and σ_{ϵ} -Mean Square Error.

The PSNR, computation time are calculated with PSO for different frame rates.

2.4. VEPSO Approach

In this approach, the number of particles is taken as 30. These particles are divided into two sub swarms of size 15. Each sub swarm is allotted to find the optimum set of subbands for their respective objective functions:

- Subswarm-1: objective function-minimum Entropy subbands
- Subswarm-2: objective function-minimum Mean Square Error Subbands

The number of iterations: 10, the parameter $c_1 = c_2 = 2$ and inertia weight is taken as 0.9. The algorithm used in this approach is same as in the previous case, the only difference is that about the best particle's information is shared by other sub swarm's particles by means of mutation and cross over operations for particle updating.

3. RESULTS AND DISCUSSION

The video sequence is first grouped into frames of 2,4 and is subjected to DDWT decomposition. After DDWT decomposition the number of coefficients is minimized using Noise Shaping algorithm. The number of coefficients is fixed as 17,000 and the multiplication factor is taken as 1.8 for the Noise shaping algorithm. The standard video sequences Foreman, Rhinos are used to test the performance of the proposed method. The average value of PSNR, MSE and Computation time for the two video sequences Foreman and Rhinos are shown in **Table 1 and 2**. The MOPSO Technique with the threshold value of 30 and number of particles 10 are given in

Table 1. Table 2 Gives the performance of the VEPSO Technique for various frame rates.

The average PSNR values are compared with the standard 3-D SPIHT algorithm as shown in **Table 3**.

For a video sequence which has many edges and motions, such like Foreman, DDWT+MOPSO outperforms SPIHT by 4 dB and VEPSO by .1 dB. For the Rhinos sequence VEPSO provides better PSNR result when compared to the other two techniques. It's PNSR performance is around 7 dB more than 3D SPIHT and 4dB more than MOPSO. The Computation time for various frame rates are both the methods are given in **Table 1 and 2**. The VEPSO computation time is higher than the MOPSO at the frame rate of 6 frames per second by 49 and 25 sec for the Forman and Rhinos sequences respectively.

Figure 3 and 4 show the variation of PSNR value for the different frames of the Foreman sequence and Rhinos video sequences. For Foreman sequence, the PSNR value increases with increasing Frame rate in the case of VEPSO, but it remains constant for MOPSO technique. In the case of Rhinos sequence the PSNR value of the reconstructed sequence is around 4 dB greater than modified technique. There is very small variation in the PSNR with variation in the number of frames for a constant threshold value.

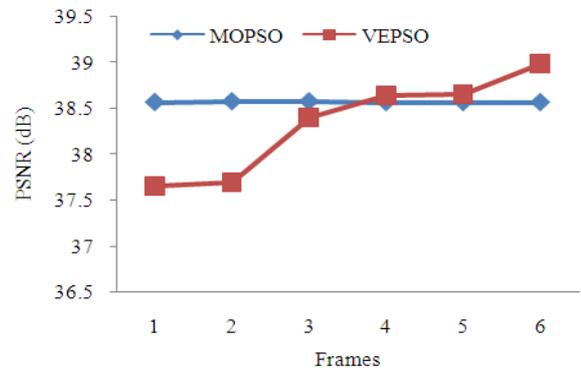


Fig. 3. PSNR plot of Foreman sequence

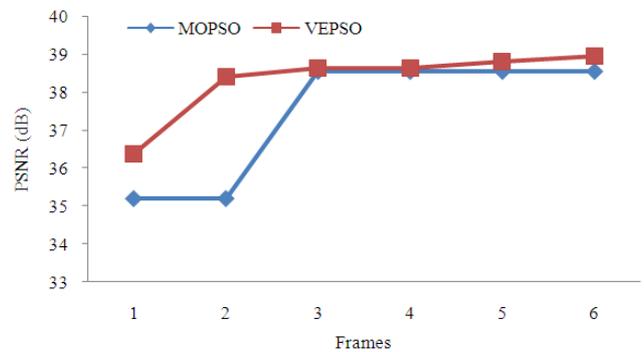


Fig. 4. PSNR plot of Rhinos sequence

Table 1. Average PSNR for video sequences (MOPSO technique)

No of frames	PSNR	MSE	PSNR	MSE	Computation	Sec
Sequence	Foreman	Sequence	Rhinos	Sequence	time Forman	Rhinos
2	38.5721	9.0338	35.1955	19.6575	27.2356	29.025600
3	38.5620	9.0547	35.1407	19.9071	34.2589	37.256800
4	38.5620	9.0548	35.1407	19.9071	41.0256	42.108900
5	38.5620	9.0548	35.1407	19.9071	47.2356	49.850200
6	38.5620	9.0548	35.1407	19.9071	105.0205	109.208102

Table 2. Average PSNR for video sequences (VEPSO Technique)

No of frames	Rhinos		Foreman Rhinos		Computation	Time Seconds
	Avg PSNR	Avg MSE	Avg PSNR	Avg MSE		
2	38.4027	9.3930	38.4027	9.3930	79.25010	80.14750
3	38.6303	8.9135	38.6373	8.9971	86.35870	87.20140
4	38.6249	8.9247	38.6489	8.8754	99.20140	100.3203
5	38.7970	8.5779	38.9855	8.2135	114.2145	115.2106
6	38.8970	8.4659	38.9855	8.2135	154.2145	134.2106

Table 3. Average PSNR for video sequences -comparison

Video sequence	3D SPIHT	Modified MOPSO	VEPSO
Foreman Sequence (at the rate of 5 Frames)	34.4563	38.5641	38.4744
Rhinos Sequence (at the rate of 5 frames)	32.4612	35.1516	39.1671

4. CONCLUSION

In this study, a video coder using 3-D dual tree wavelet transform combined with modified Multi objective particle swarm optimization and Vector Evaluated PSO are proposed and tested. The multi objective functions of MSE and Entropy are considered for both approaches. The modified MOPSO method is simple and less complex when compared to the VEP SO approach. The proposed methods are tested on standard video sequences. The test results are compared with the standard 3D SPIHT coding method. The obtained results using VEP SO technique show that the PSNR value is comparatively good with MOPSO-Weighted aggregate method. But the computation time is more when compared to the 3D SPIHT and Weighted aggregate MOPSO method. By considering the parallel processing of VEP SO, it is more suitable for hardware implementation and also the quality of the reconstructed video sequence is reasonable (above 38 dB). In future, the variants of the PSO and also other approaches for Multi objective optimization will be considered in order to improve the performance of the video coding system with lesser complexity.

5. REFERENCES

- Boukhobza, A., A.T. Ahmed, N. Taleb and A. Bounoua, 2013. Optimization design of orthogonal filter banks for image coding via multi-objective genetic algorithm. Department of Electronics, UHBC University.
- Chin-Wei, B. and M. Rajeswari, 2010. Multiobjective optimization approaches in image segmentation-the directions and challenges. *Int. J. Adv. Soft Comput. Applic.*, 2: 40-65.
- Coello, C.A.C., C.C. Coello and M.S. Lechuga, 2004. Handling multiple objectives with particle swarm optimization. *IEEE Trans. Evol. Comput.*, 8: 256-279. DOI: 10.1109/TEVC.2004.826067
- Konak, A., D.W. Coit and A.E. Smith, 2006. Multi-objective optimization using genetic algorithms: A tutorial. *Reliab. Eng. Syst. Safety*, 91: 992-1007. DOI: 10.1016/j.res.2005.11.018
- Ouyang, C., X. Li and N. Wang, 2010. A best wavelet packet basis image compression algorithm based on PSO. *Proceedings of the 4th International Conference on Genetic and Evolutionary Computing*, Dec. 13-15, IEEE Xplore Press, Shenzhen, pp: 11-13. DOI: 10.1109/ICGEC.2010.11
- Parsopoulos, K.U. and M.N. Varahatis, 2002. Particle swarm optimization method in multiobjective problems. *Proceedings of the ACM Symposium on Applied Computing*, Mar. 10-14, ACM Press, New York, USA., pp: 603-607. DOI: 10.1145/508791.508907
- Thamarai, M. and R. Shanmugalakshmi, 2010. Video coding technique using 3-D dual tree complex wavelet Transform with improved particle swarm optimization. *IACSIT Int. J. Eng. Technol.*, 2: 296-301.
- Thamarai, M. and R. Shanmugalakshmi, 2011. *Swarm Intelligence in Video Coding*. In: *Recent Advances on Video Coding*, Lorente, J.D.S. (Ed.), Intech Open Access Publications.
- Yang, J., B. Wang, Y. Wang and W. Xu, 2007. Video Coding using 3-D anisotropic dual-tree wavelet transform. *Proceedings of the IEEE International Conference on Multimedia and Expo*, Jul. 2-5, IEEE Xplore Press, Beijing, pp: 1371-1374. DOI: 10.1109/ICME.2007.4284914