

Color Image Compression Based On Adaptive Spectral Transform and Coded Decision

S.Vashni Abiya¹, Seebika², B.Kathiravan³ & A.Akbar Ali⁴

UG Student, Department of ECE, Rajas Engineering College, Vadakkangulam, India^{1,2,3,4},

ABSTRACT: The different texture of a color image generally exhibit different frequency characteristics, the energy condense of applying a single spectral transform to all texture is largely inefficient from a compression prospect. Thus, it is proposed that different subsets of wavelet coefficients of a color image be subjected to different spectral transforms before the resultant coefficients are coded by an efficient wavelet coefficient coding scheme such as that used in Color Set Partitioning In Hierarchical Trees (CSPIHT). A quad-tree represents the spatial partitioning of the set of high frequency coefficients of the color planes into spatially oriented subsets which may be further subdivided into smaller directionally oriented subsets. The partitioning decisions and Rate-Distortion (R-D) optimized by employing a known analytical R-D model.

KEYWORDS: Adaptive coding, color image coding, Karhunen–Loeve transform, quad-tree, wavelet transform, CSPIHT, Rate Distortion, color planes, spatial transform, 2D-DWT.

1.INTRODUCTION

The choice of the compression method, lossy or lossless, depends on the application. For example, in applications dealing with speech signals and video television images, where some loss of information can be tolerated, lossy compression method can be used. On the other hand in a wide range of medical applications and under special circumstances such as disease diagnostic, the loss of information is unacceptable.

Wavelet analysis is an exciting new method for solving difficult problems in mathematics, physics, and engineering, with modern applications as diverse as wave propagation, data compression, signal processing, image processing, pattern recognition, computer graphics, the detection of aircraft and submarines and other medical image technology. Wavelets allow complex information such as music, speech, images and patterns to be decomposed into elementary forms at different positions and scales and subsequently reconstructed with high precision. Signal transmission is based on transmission of a series of numbers. The series representation of a function is important in all types of signal transmission. The wavelet representation of a function is a new technique.

Wavelet transform allows the components of a non-stationary signal to be analyzed. The CSPIHT algorithm links the SOTs for the three spectral planes of coefficients by defining the nodes in the lowest frequency sub-bands of the chrominance planes second and third eigenvector component planes for Karhunen–Loeve Transform (KLT) as the children of the spatially corresponding nodes in the lowest frequency sub-band of the luminance plane (principal eigenvector component plane for KLT). Due to the structure of the overall tree, only luminance nodes and sets rooted at them are tested for significance at the initial passes. This is unlike CEZW where the chrominance root nodes or sets rooted at them are unnecessarily separately tested for significance at the initial passes even though they are almost always insignificant. Thereby, a savings in bit expenditure is achieved with respect to CEZW at low rates. However, for a discrete wavelet transform with a large number of levels, the performance gain of CSPIHT is significant only at very low rates. At higher rates, the gain is negligible since the difference between the number of SOT roots for CSPIHT and the number of SOT roots for the benchmark method is small and the SOT roots are only visited in the bit planes visited early in the coding process. The SOT of CSPIHT was later extended to link spatially related blocks of coefficients. By revealing that DWT yields correlations between luminance nodes in the spatial vicinity of each chrominance node.

International Journal of Innovative Research in Science, Engineering and Technology

An ISO 3297: 2007 Certified Organization

Volume 4, Special Issue 11, September 2015

National Conference on Emerging Trends in Computing, Communication & Control Engineering [NCET 2K15]

On 4th September, 2015

Organized by

Department of CSE, ECE & EEE, Magna College of Engineering, Chennai-600055, India.

Vector SPIHT is an extension of SPIHT that builds a SOT on vectors of DWT coefficients at corresponding locations in the three color planes to exploit the statistical dependencies among them. For video coding, decent gains over H.263 at medium to high rates.

In a region adaptive 1-D KLT across the color planes is followed by the spatial 2-D discrete cosine transform (DCT) within each spectral plane. A quad-tree represents the recursive subdivision of the image is divided into variable size blocks. Each block is transformed with the KLT derived from its covariance matrix. The quad-tree and the quantizer for the transform coefficients are jointly R-D optimized.

The primary contribution of the current work is the specification of a R-D optimized design procedure for an adaptive spectral transform whose color image compression performance surpasses that of the widely used single global spectral transform at all rates when integrated with zerotree or block-based DWT coefficient coding algorithms. In its minimization of a R-D cost function, the proposed method is somewhat similar to the coding mode selection process formulated for video coding. Coding mode selection has been extended to the adaptive selection of a fixed spectral transform from a small set of candidate, fixed spectral transforms for each fixed size block of the frame residual prediction error in a H.264/AVC advanced 4:4:4 profile setting. Unlike the proposed method, the side information rate for conveying the spectral transform identity is very little and not factored into the cost functions of the candidate, fixed transforms

II. FEATURES OF THE ADAPTIVE TRANSFORM

The decorrelating transforms such as the signal dependent KLT or the fixed RGB-YCrCb are commonly applied to the color planes of a color image prior to encoding so as to compact most of the signal energy in the three color planes to a single spectral plane (the principal eigenvector component plane for KLT and the Y component plane for YCrCb) and thereby realize a compression performance advantage. At the end of decoding, a corresponding inverse spectral transform yields the reconstructed color planes.

A. ORDER OF THE SPATIAL AND SPECTRAL TRANSFORMS

The computationally efficient 2-D DCT, the spatial transform employed in the current paper is the performance efficient 2-D DWT that is widely adopted in the state-of-the-art image coding. As shown in Fig. 1, the spatial 2-D DWT transform is not preceded, but followed by one or more spectral 1-D KLT transform(s), each of which is uniquely determined for a subset [region(s) or block(s)] of a partition of DWT coefficients of the image. If the spatial 2-D DWT were to follow the 1-D spectral KLT transform(s) of image pixels, the sub-band analysis filter kernels would have to be applied across boundaries between two subsets of spectral coefficients obtained with different KLT bases. Filtering across boundaries may be circumvented by mirror extending the filtered values at the boundaries. However, performing a mirror extension at the boundary between each pair of subsets significantly reduces the performance efficiency of 2-D spatial DWT. spectrally transformed arbitrary regions of the image.

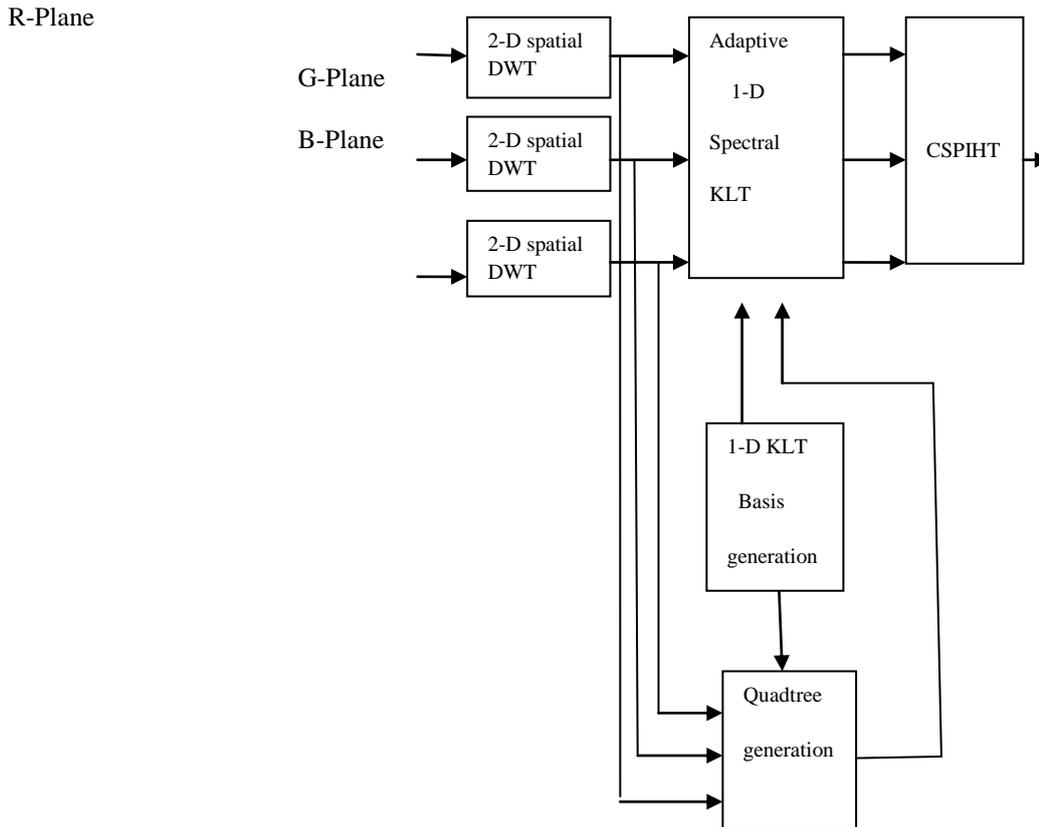


Fig 1. Block diagram of proposed system.

A second reason for the order of transforms is that, in the DWT domain, it is possible to further partition a set of spatially oriented coefficients (corresponding to a region or a block in the image) into smaller subsets of spatially as well as directionally oriented coefficients. This way, the KLT bases may be better adapted to the edge directionality. A final reason is that, in each color plane, each set of spatially and/or directionally oriented high frequency DWT coefficients has near zero mean value. Such mean values do not need to be coded as side information for the application of the 1-D KLT across the color planes of DWT coefficients. Had the 1-D KLT been performed directly on the input image prior to 2-D DWT, the subtraction and subsequent coding of the mean value for each set of image pixels in each color plane would have been necessary.

B. STRUCTURE OF THE ADAPTIVE SPECTRAL TRANSFORM AND CODED DECISIONS

A precise description of arbitrary shaped sets of coefficients of uniform content is expensive from a compression viewpoint even at high coding rates. The viable solution proposed is also employed in other contexts such as variable block size motion estimation, is to spatially partition the image into variable size blocks of near uniform content. In the current paper, the layout of the subsets of coefficients is concisely represented by a quad-tree structure.

When a set of coefficients represented by its KLT basis is partitioned into smaller subsets represented by their own KLT bases, the energy compaction of the resultant coefficients improves, but the R-D performance may or may not improve due to the increase in side information for coding the increased number of KLT bases. This suggests that one should decide to split a quad-tree node by partitioning a set or, conversely, prune a sub-tree of the quad-tree by merging small subsets into a large set only if a R-D improvement is expected as a consequence. Since the spectral 1-D KLT is

applied on the 2-D DWT coefficients, it is conceivable that, for each sub-band, the spatial partition be unique and represented by a different quad-tree structure. However, such an approach necessitates a substantial side information rate at low rates. In order to conserve this rate, the strong statistical dependencies among the magnitudes of coefficients of the same spatial orientation across high frequency sub-bands may be exploited. All high frequency subbands are constrained to be spatially partitioned into blocks of coefficients in the same way.

For each set of spatially oriented blocks of coefficients across all high frequency sub-bands, a second decision is made as to whether or not to partition it three directionally (vertically, horizontally, and diagonally) oriented subsets. A unique basis for each of these three subsets can represent and exploit the statistical dependencies among the color components within that subset in a better way. It is favored if the R-D cost is reduced as a consequence. To improve the coding performance, a final decision is made as to whether to project each subset of coefficients onto the fixed YCrCb basis instead of the signal dependent KLT basis. The fixed basis is favored over the KLT basis if the R-D cost of the fixed basis is smaller due to excessive side information for coding the KLT basis vectors.

C. CODING OF THE COEFFICIENT WITH CSPIHT

Both the spatial 2-D DWT and the spectral 1-D KLT compact the signal energy to a small number of transform coefficients. Superior coding performance can be achieved if this non-uniform distribution of the signal energy is exploited by an efficient bit allocation strategy. Toward this end, one may employ a set partitioning-based wavelet coefficient coding scheme for color images large magnitude coefficients early in the coding process and explicitly or implicitly allocates rate to the spectral as well as the spatial components.

The well-known grayscale image coding method SPIHT encodes and transmits wavelet coefficients in multiple coding passes corresponding to bit planes (in the order from the most significant to the least significant). In each coding pass, only the wavelet coefficients with magnitudes exceeding a certain threshold (significant coefficients) are encoded. The coefficients are ordered in hierarchies, called SOTs, with roots in the lowest frequency sub-band, branching successively into higher frequency sub-bands at the same spatial orientation. The most significant bit of a significant coefficient is encoded by identifying its location via recursively subdividing the spatial orientation tree into smaller sub-trees and single coefficient sets. In subsequent passes, the refinement bits (lesser significant bits) of a significant coefficient are also coded by successive approximation quantization. The algorithm has high performance due to representation of large sets of zero bits of coefficients by a single symbol. Since the performance of the benchmark method is generally inferior to that of CSPIHT for DWT with few levels or for very low coding rates, CSPIHT has been adopted for jointly coding the coefficients of the three spectral planes.

III. R-D OPTIMIZED ADAPTIVE TRANSFORM DESIGN

Let the rate equivalent of the R-D cost estimate at the R-D operating point (R, D) be defined as $J^\wedge = D^\wedge / S^\wedge + R^\wedge$ where $D^\wedge, R^\wedge, S^\wedge$ are estimates of the distortion, rate, and the absolute value of the slope at this point, respectively. This section first treats how these estimates are obtained for wavelet-based coding in CSPIHT.

A. Distortion, Rate, And Slope Estimates

The expected distortion of the successive approximation quantizer with a threshold (step size) of Δ can be expressed as a weighted sum of the conditional expected distortions for the insignificant and significant coefficients as follows:

$$D_{coef}(\Delta) = E[|X - X^\wedge|^2 \mid |X| < \Delta]pr\{|X| < \Delta\} + E[|X - X^\wedge|^2 \mid |X| \geq \Delta]pr\{|X| \geq \Delta\} \dots\dots\dots(1)$$

Where X and X[^] are the original and reconstructed values.

The distortion estimate is formed as follows

$$D_{coef}^\wedge(M_n) = \frac{1}{n} (\sum_{i=M_n+1}^n |x(i)|^2) + \frac{M_n}{4\lambda^2} h(K, s) \dots\dots\dots(2)$$

where $k=\lambda\Delta$, $x(i)$ is the coefficient of the i th largest magnitude, λ is the parameter of the Laplacian distribution estimated as the inverse of the mean magnitude of a coefficient, and

$$h(k,s) = \frac{1 - e^{-\frac{ks}{2}}}{e^k - 1} (k^2 - 4k + 8 - e^{-k}(k^2 + 4k + 8)) \dots \dots \dots (3)$$

Prior to the actual encoding, the distortion may also be determined at the coding pass boundaries by a computationally more expensive, but also more exact alternative to the estimate provided by (2). Specifically, the quantization strategy of the specific wavelet coefficient coding scheme is simulated. On the contrary, the estimate for the rate of the DWT coefficients prior to their coding, given as

$$R_{coef}^{\wedge}(M_n) = \frac{c}{n} M_n \dots \dots \dots (4)$$

is only approximate. Even though the assumption of linearity of the dependence of rate on Mn is valid for most images, the linearity constant C is in a wide range of 3–7.

The total rate estimate for coding a total of N coefficients in the entire image can be expressed as

$$R_{tot}^{\wedge} = R_{coef}^{\wedge}(M_N) + R_{sideinfo} \dots \dots \dots (5)$$

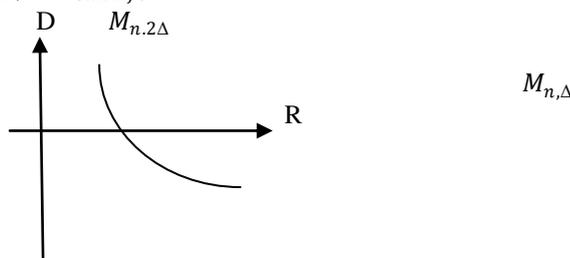


Fig 2. Distortion-rate characteristic two adjacent coding passes

Side information rate, Rsideinfo, consists of the number of bits necessary to convey to the decoder the quad-tree structure (1 bit for each internal node of the quad-tree), one flag bit for each subset of the final partition that signals whether a fixed or a signal dependent KLT basis is used, and the bits representing the components of the KLT basis vectors. For each KLT basis, only two components of two of the basis vectors are coded by representing them in a polar coordinate system. The radial component is coded with 12 bits precision and the angular component is coded with a precision that increases linearly from 6 to 8 bits with radius. The third component of these vectors is derived by applying the unit norm property. The third vector is derived as the cross product of the other two vectors.

The absolute value of the slope $S = |dD_{coef} / dR_{coef}|$ is estimated as

$$S^{\wedge}(M_n) = \frac{D_{coef}^{\wedge}(M_n) - D_{coef}^{\wedge}(M_{n-1})}{R_{coef}^{\wedge}(M_{n-1}) - R_{coef}^{\wedge}(M_n)} \dots \dots \dots (6)$$

IV. PROPOSED ENCODING METHODOLOGY

A new scheme, that limits the complexity to two coding iterations, is proposed in this paper for CSPIHT-based encoding. For each $i = 0, 1, 2, \dots, I$ let $\Delta_i = \Delta_0 2^{-i}$. The steps of proposed scheme are as follows.

- 1) Initial step.
 - a) The target rate R^{tar} is specified by the user.
 - b) All 2-D DWT coefficients of the color planes are spectrally transformed by a single KLT basis. Let $R_{coef}^{tar} = R^{tar} - R_{sideinfo}$.
 - c) For each $i = 0, 1, 2, \dots, I$: M_{N,Δ_i} number of significant coefficients that exceed Δ_i are determined to yield the rate estimate $R_{coef}^{\wedge}(M_{N,\Delta_i})$ by using (4).
 - d) Δ_p^{\wedge} is initialized: $\Delta_p^{\wedge} = \max_{\{i: R_{coef}^{\wedge}(M_{N,\Delta_i}) R_{coef}^{tar}\}} \Delta_i$.
- 2) For each of two coding iterations.
 - a) The adaptive transform is designed by employing rate and slope estimates based on the previously determined Δ_p^{\wedge} and applied to spectrally transform the 2-D DWT coefficients. R^{tar} is updated: $R_{coef}^{tar} = R^{tar} - R_{sideinfo}$.
 - b) The coefficients are CSPIHT encoded at rate $R_{coef}(\Delta_i)$ in the first and R_{coef}^{tar} in the second iteration.

c) For each $i=0,1,2,\dots,I$, the rate $R_{coef}(\Delta_i)$ at the end of the pass for threshold Δ_i is determined.

d) Δ_p is updated: $\Delta_p = \max_{\{i:R_{coef}(M_{N,\Delta_i})R_{coef}^{tar}\}} \Delta_i$.

The bit stream generated in the second CSPIHT encoding of the coefficients is the actual transmitted bit stream. Experiments on all the test images at all the rates

Rate(b/p)	Average PSNR(dB) of Y plane,UV planes					
	Single spatial domain KLT basis+CSPIHT	Single DWT domain KLT Basis+CSPIHT	Singles spatial domain KLT basis+JPEG2000	Two DWT domain KLT Basis+CSPIHT	Adaptive transform+CSPIHT	
					Sim.ofquantization	Distortion estimate
0.05	27.03, 37.19	27.02, 37.20	26.96, 37.19	27.01, 37.22	27.03, 37.31	27.01, 37.32
0.10	28.93, 38.67	28.92, 38.67	28.96, 38.79	28.93, 38.76	28.96, 37.31	28.94, 37.32
0.25	32.01, 40.76	32.01, 40.79	31.98, 40.93	32.04, 41.01	32.09, 41.26	32.08, 41.23
0.50	35.02, 42.66	35.02, 42.70	34.99, 42.82	35.09, 43.08	35.17, 43.35	35.14, 43.36

Table1:Performances(average over 32 images) of the proposed and reference methods at various rates.

tested have shown that Δ_p convergences in two coding iterations to the Δ_p value given by the trivial exhaustive search scheme.

In the first coding iteration of the proposed scheme, the estimates are obtained with (4) and (5) as before. In the second iteration, the estimates are optimized by using data collected in the first iteration. Let p_Δ and $1-p_\Delta$ be the probabilities of coding and not coding, respectively, a coefficient in the final coding pass at rate R_{coef}^{tar} . The value of p_Δ is estimated as the fraction of the number of coefficients coded in the final coding pass of the first coding iteration. For this purpose, each coefficient coded in this pass is tagged to be counted later on.

In the second coding iteration, the proportionality constant, chosen in an ad hoc manner is adjusted adaptively as follows:

$$C = \frac{R^1 - R_{sideinfo}^1}{(p_\Delta M_{N,\Delta}^1 + (1-p_\Delta)M_{N,2\Delta}^1)/N} \dots\dots\dots(7)$$

In this equation, R^1 and $R_{sideinfo}^1$ are the rate and side information rate, respectively, and $M_{N,\Delta}^1$ and $M_{N,2\Delta}^1$ are the number of significant coefficients in the entire image at the end of the final pass and at the end of the pass before the final pass, respectively, in the first coding iteration.

V. EXPERIMENTAL RESULTS AND DISCUSSION

In this section, we report and discuss results of simulations conducted to compare the compression performance of compression systems that integrate the adaptive transform and the coefficient coding schemes of CSPIHT against the performances of several reference compression systems.

A large set of 32 color images has been used to evaluate the compression performances. Twenty-four of the images are the 768×512 or 512×768 Kodak true color 24 bit images, six of them are the 512×512 24 bit images, Lena, Baboon, Tiffany, Sailboat, Airplane, and Pepper, and the other two are the high resolution 2048×3072 24 bit images artificial. Throughout the simulations, the DWT filters used are the CDF 9/7 biorthogonal filters.

Let d_γ be the mean squared error of the color or spectral plane γ . By letting $d^- = d_\gamma$, $d^- = \frac{d_u + d_v}{2}$ and $d^- = (d_R + d_G + d_B)/3$, the peak signal-to-noise ratios (PSNRs) of the Y component plane, UV component planes, and the color image are evaluated as

$$PSNR_{dB} = 10 \log_{10}((255)^2/d^-) \dots \dots \dots (8).$$

The performance gain of the “adaptive transform” over the “single spatial domain KLT basis” increases with rate upto as much as an average of 0.9 dB in the Y component and an average of 1.35 dB in the U and V components at 2.5 b/p. In this case, the large side information rate due to the transmission of a large number of bases vectors for a complex partition is more easily afforded at high rates. On the contrary, the average gain of the “adaptive transform” over the “two DWT domain KLT bases” is limited to 0.25 dB in Y component and 0.4 dB in U and V components at medium to high rates since not all images have local regions of distinct color transitions that can be exploited with the allocation of unique bases to these regions. For example, in Kodak color images number 5 and number 8, the sharp color transitions of contours of motorbikes and roofs occur in close proximity and similar high frequency content occurs over most of these images which cannot be isolated distinctively in local regions. Moreover, even when such an isolated local region of distinct color transition does exist in an image, it need not coincide with a block structured subset of the finest quad-tree partition. The highest gains with the “adaptive transform” have been observed for the artificial image which has sharp color transitions along edges of artificial objects that can be isolated in local regions.

V. CONCLUSION

This paper has demonstrated that a locally adaptive color spectral transform applied in the DWT domain across the color planes of an image yielded a compression performance advantage that increases progressively with rate upto an average of 0.9 dB in the Y and 1.35 dB in the U and V components over a single global transform applied in the spatial domain. This advantage is largely due to the application of separate bases to the low and high frequency DWT coefficients. This subset of the proposed method increased complexity only slightly. The proposed method also employed quad-tree-based partitioning of high frequency coefficients into spatially and/or directionally oriented subsets where the decisions to form the partition as well as to employ fixed or signal dependent basis within each subset are RD optimized by the distortion and rate models. The application of a separate basis to each subset is a fine-tuning for further enhancing the performance in return for a considerable increase in computational complexity. The complexity of the encoder that precisely achieved a given target rate can be limited with two iterations of adaptive transform design and subsequent CSPIHT coefficient coding. The proposed method yielded limited improvement when extended to be applied to multispectral/hyperspectral image data, primarily due to the linear increase of side information rate with the number of coded bands. In the future work, a hierarchical application of the KLT transform for lowering the side information rate for this kind of data will be investigated.

REFERENCES

- [1] Cagnazzo M, Gaetano R, Parrilli S, and Verdoliva L, “Region based compression of multispectral images by classified KLT,” in *Proc. 14th EUSIPCO*, Sep. 2006.
- [2] Deforges O, Babel M, Bedat L, and Ronsin J, “Color LAR codec: A color image representation and compression scheme based on local resolution adjustment and self-extracting region representation,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no. 8, pp. 974–987, Aug. 2007
- [3] Gelli G and Poggi G, “Compression of multispectral images by spectral classification and transform coding,” *IEEE Trans. Image Process.*, vol. 8, no. 4, pp. 476–489, Apr. 1999.
- [4] Kassim A and Lee W.S, “Embedded color image coding using SPIHT With partially linked spatial orientation trees,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 2, pp. 203–206, Feb. 2003
- [5] Marpe D, Kirchhoffer H, George V, Kauff P, and Wiegand T, “An adaptive color transform approach and its application in 4:4:4 video coding,” in *Proc. 14th EUSIPCO*, Sep. 2006.

International Journal of Innovative Research in Science, Engineering and Technology

An ISO 3297: 2007 Certified Organization

Volume 4, Special Issue 11, September 2015

National Conference on Emerging Trends in Computing, Communication & Control Engineering [NCET 2K15]

On 4th September, 2015

Organized by

Department of CSE, ECE & EEE, Magna College of Engineering, Chennai-600055, India.

- [6] Pearlman W.A , Islam A, Nagaraj N, and Said A, Efficient, low-complexity image coding with a set-partitioning embedded block coder,"*IEEE Trans. Circuits Syst. Video Technol.* , vol. 14, no. 11, pp. 1219–1235, Nov. 2004
- [7] Mallat S and Falzon F, "Analysis of low bit rate image transform coding," *IEEE Trans. Signal Process.*, vol. 46, no. 4, pp. 1027–1042, Apr. 1998.
- [8]. San X, Cai H, and Li J, "Color image coding by using inter-color correlation, in Proc. *IEEE Int. Conf. Image Process.*, Oct. 2006, pp.3117-3120.
- [9]. Shen K and Delp E.J, "Color image compression using an embedded rate scalable approach," in Proc. *IEEE Int. Conf. Image Process.*, Oct.1997,pp.34-37.
- [10]. Wiegand T, Lightstone M, Mukherjee D, Campbell T.G, and Mitra S.K, "Rate distortion optimized mode selection for very low bit rate video coding and the emerging H.263 standard," *IEEE Trans. Circuits Syst.Video Technol.*, vol. 6, no. 2, pp. 182–190, Apr. 1996.