

Intra Prediction Mode Decision for H.264

Mamatha R.B.
Research Scholar
Dept of ECE, EPCET
Bengaluru, Karnataka, India

Keshaveni N., PhD
Professor
Dept of ECE, KVGCE
Sullia, Karnataka, India

ABSTRACT

Digital video compression has become an integral part of the way we create, consume visual information and communicate, over the last few decades. A robust technique of compression is proposed in this paper for video compression. Reduction of irrelevant or redundant data in order to save the storage space requirements and processing time is nothing but compression. Here H.264/AVC (Advance Video Coding) video coding standard is used for compression. Where I-frames are divided into different macro blocks (MB) and each MB is efficiently compressed using 4 x4 and 16 x16 blocks in H.264/AVC intra prediction. Choosing one of 9 prediction modes for each 4 x4 block with reduced time and less complexity is still a bottleneck. In this paper a gradient based fast intra prediction mode selection method for 4x4 and 16x16 is proposed. The proposed method divides an input frames into variable block sizes based on the texture and performs few mode examinations based on the gradient direction in the given slice. Respective best mode with minimum cost is selected using sum of absolute difference (SAD) and rate distortion optimization (RDO). This process is carried for all the video input frames. Each compressed video frames are combined finally to get a compressed video output.

Keywords

Advanced Video Coding (AVC), compression, H.264, SAD, RDO, MB and Gradient method.

1. INTRODUCTION

H.264 is the latest video compression standard for advanced video (AVC), in the coming years it is expected to become the video standard of choice. H.264 supports the most efficient video compression available today. Compared to other format H.264 encoder can reduce the size of a digital video file by more than 80%. At high and low bit rates also it offers good video quality.

H.264 include 4x4 integer DCT, intra prediction in I-frame, quarter-pixel motion compensation, multiple reference frames and multiple block size for P-frame etc type of advanced features. These features help H.264 to achieve higher efficiency and compression significantly than the existing standards. To achieve such high efficiency, complexity and computational load of the codec increased drastically compared to other codec. Thus there is a great need of high efficient algorithms to overcome this problem. To choose the best MB mode, H.264 encoder calculates the RDcost (Rate distortion cost) of every possible mode and chooses the mode having the minimum bit cost. Due to which the computational complexity is increased drastically compared with previous standards, so it makes difficult for H.264/AVC applications with low computational capability, such as mobile devices. Complexity reduction for H.264 has been an active research area. The main objective is to reduce the complexity while

keeping RD performance as close to that of full search as possible.

H.264 video encoder carries out prediction, transform and encoding processes to produce a compressed H.264 bit stream. The prediction of macro block is done based on previously coded data either from current frame (intra prediction) or from other frames which have been coded and transmitted (inter prediction). Finding a suitable inter prediction is often described as motion estimation. Subtracting an inter prediction from the current macro block is motion compensation. An n obtained residual block sample is transformed using 4x4 or 8x8 integer transform. The transformed output is quantized by dividing it by any integer value. According to the quantization parameter obtained prediction of the transform coefficients is reduced. To form the compressed bit stream video coding process produces a number of values that must be encoded. Considering the RDO procedure for intra mode selection in H.264/AVC, the number of mode combinations in one macro block is $N8 \times (16 \times N4 + N16) = 8 \times (16 + 16) = 592$. where $N8$ indicates the number of modes of an 8x8 chroma block, $N4$ indicates number of modes of an 4x4 luma block, $N16$ indicates number of modes of an 16x16 luma block. The H.264/AVC encoder carries out 592 RDO calculations to choose the best matching MB. As a result, the complexity of the encoder increases extremely.

A. Elyousfi, A. Tamtaoui and E. Bouyakhf [3] proposed a method which uses H.264/AVC standard intra prediction, 9 directional modes for 4x4 and 16x16 intra blocks. For each MB RDO calculation is reduced to less than 50 from 592 using quadratic and gradient model for homogeneous and non homogeneous block. Prediction of macroblock is done based on previously coded data either from inter

Mohammed Golam Sarwer and Q. M. Jonathan Wu [5] proposed a method, in which the number of prediction modes for each 4x4 blocks is selected adaptively. Each block is classified as one of three categories, based on similarities of the reference pixel. Video quality is improved and bit rate is saved using this method.

Bo Shen, Kuo-Hsiang Cheng, Yun Liu and Ying-Hong Wang [8] introduced a method to reduce the computational complexity of H.264/AVC High profile. Two methods are combined in this algorithm. The first method involves a quant-based Block-size selection decision, which is sum of the quantization based with error adjustment to select either intra 4 x 4 or intra 16 x 16 mode predictions. The second method involves direction based prediction mode decision through which they could reduce the possible prediction modes for the rate-distortion (RD) optimization technique.

Chitradevi s and Vijayalakshmi K [9] proposed a technique of compression which used H.264/AVC (Advance Video Coding) video coding standard. Reduncy of image data is reduced using intra prediction technique during encoding phase. This is

performed in the current frame not relative to any other frame. In addition to improve the encoding efficiency the integer transform and the CABAC contribute. For improving encoding efficiency project can be further enhanced.

Xiaohong Zhang [10] proposed a technique which adopts H.264/AVC RDO approach to search all the prediction modes to determine the best mode with less bit rate transcoding. Incompatibility issues of video stream transmission are solved using Frame-skipping transcoding among heterogeneous networks.

Robust method using typical block type intra 4 x 4 and intra 16 x 16 blocks are discussed here in this paper to reduce the number of RDO calculations. Here variable block sizing is done based on the texture feature present in the MB. Mode selection is done using gradient calculation. Gradient calculation is applied for every slice in an input frame to select the modes. Best MB Mode with maximum visual quality and minimum bit cost is calculated using SAD and RDO, which is explained briefly in the following section.

2. PROPOSED WORK

The working flow in our proposed methodology is as illustrated in Figure 1. Video inputs are taken and frames are generated from the input video. Each frame undergoes pre-processing which includes resizing and normalization of frames. Next step is to find I frame using group of pictures (GOP).

2.1 GOP

The order in which the intra and inter frames are arranged are specified by GOP or group of pictures. GOP is nothing but a group of successive pictures. A GOP consists of I picture or I frames (intra coded picture) .I frames are coded independently of all other pictures. An I frames indicates the beginning of each frames. I frames are always followed by P and B frames. I frames always contain the full image and additional information is not needed to reconstruct it.

The GOP structure always contains two numbers, for example P, Q. First number indicates the distance between two anchor frames i.e. I and P. Second tells the difference between the two I frames. For example if we take P=3, Q=12, the obtained GOP structure is IBBPBBPBBPBB. After the I frame selection, each input I frame is divided into slices. Each input slice is divided into N no of macro blocks (MB). Each MB is divided into different blocks and size of each block is obtained based on the texture feature present in the MB. If the input MB contains more homogeneity then it is divided into 16x16 blocks and if the homogeneity is less, then the MB is divided into 4x4 blocks. Usually the MB size will be of 16 x16. Gabor wavelet transform is done for texture feature extraction.

2.2 Gabor wavelet transform

Space and spatial frequency bounded by the relations $\Delta_x^2 \cdot \Delta_u^2 \geq \frac{1}{4\pi}$ and $\Delta_y^2 \cdot \Delta_v^2 \geq \frac{1}{4\pi}$ where $[\Delta_x^2, \Delta_u^2]$ Gives spatial resolution in space and $[\Delta_y^2, \Delta_v^2]$ gives spatial frequency resolution. Good performance in texture discrimination and segmentation is given by Gabor filter. 2D Gabor filter, produce a strong correlation with an actual human segmentation during texture analyzers implementation. Gabor functions are Gaussians modulated by complex sinusoids. The two dimension equation is given by:

$$g(x,y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left(-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right) + 2\pi jwx\right) \quad (1)$$

Gabor filter masks are scale and orientation tunable, line and edge detectors. Underlying texture information is characterized by statistics of these micro features in a given region. Class of such self similar functions is referred to as Gabor wavelets. By appropriate rotation and dilation of $g(x,y)$, self-similar filter dictionary can be obtained through the generating function. Later gradient calculation is done in the begging of every slice for first MB. Only few modes are selected based on the gradient direction obtained.

2.3 Gradient calculation

Partial derivative of the smoothly varying inhomogeneity is calculated by gradient calculation. An inhomogeneity multiplicative model in which acquired image $v(x)$ at a voxel x , which is given by:

$$V(x) = g(x)u(x) + n(x) \quad (2)$$

where $g(x)$ is the inhomogeneity gradient with spatial varying, $u(x)$ is the ideal signal and $n(x)$ is the additive noise. The difference in intensity between the pixel at (x, y) and adjacent pixel along the x direction is defined as:

$$\begin{aligned} \Delta_x v(x,y) &= v(x+1,y) - v(x,y) \\ &= \Delta_x g(x,y) \cdot u(x,y) + g(x+1,y) \cdot \Delta_x u(x,y) \\ &\quad + \Delta_x n(x,y) \end{aligned} \quad (3)$$

The first, second and third are not necessary small compared to the inhomogeneity term, unless computation is performed on a high SNR and homogeneous region. Second term indicates the inherent intensity variation arising from the biological tissue properties and third term indicates the noise. To minimize the effect of these two terms during calculation in an imperfect environment, we make use of set of n differences between adjacent pixels along y direction at the same. The sum of these is:

$$\begin{aligned} \sum_{j=0}^{n-1} \Delta_x v(x_0, y_0 + j) &= \sum_{j=0}^{n-1} \Delta_x g(x_0, y_0 + j) \cdot v(x_0, y_0 \\ &\quad + j) \\ &\quad + \sum_{j=0}^{n-1} g(x_0 + 1, y_0 + j) \cdot \Delta_x u(x_0, y_0 + j) \\ &\quad + \sum_{j=0}^{n-1} \Delta_x n(x_0, y_0 + j) \end{aligned} \quad (4)$$

It is shown that if n is very large, and then second and third terms can be neglected in Eq (4).

$$\begin{aligned} \sum_{j=0}^{n-1} \Delta_x v(x_0, y_0 + j) &\cong \Delta_x g(x_0, y_{av}) \\ &\quad + j) \cdot \sum_{j=0}^{n-1} v(x_0, y_0 + j) \end{aligned} \quad (5)$$

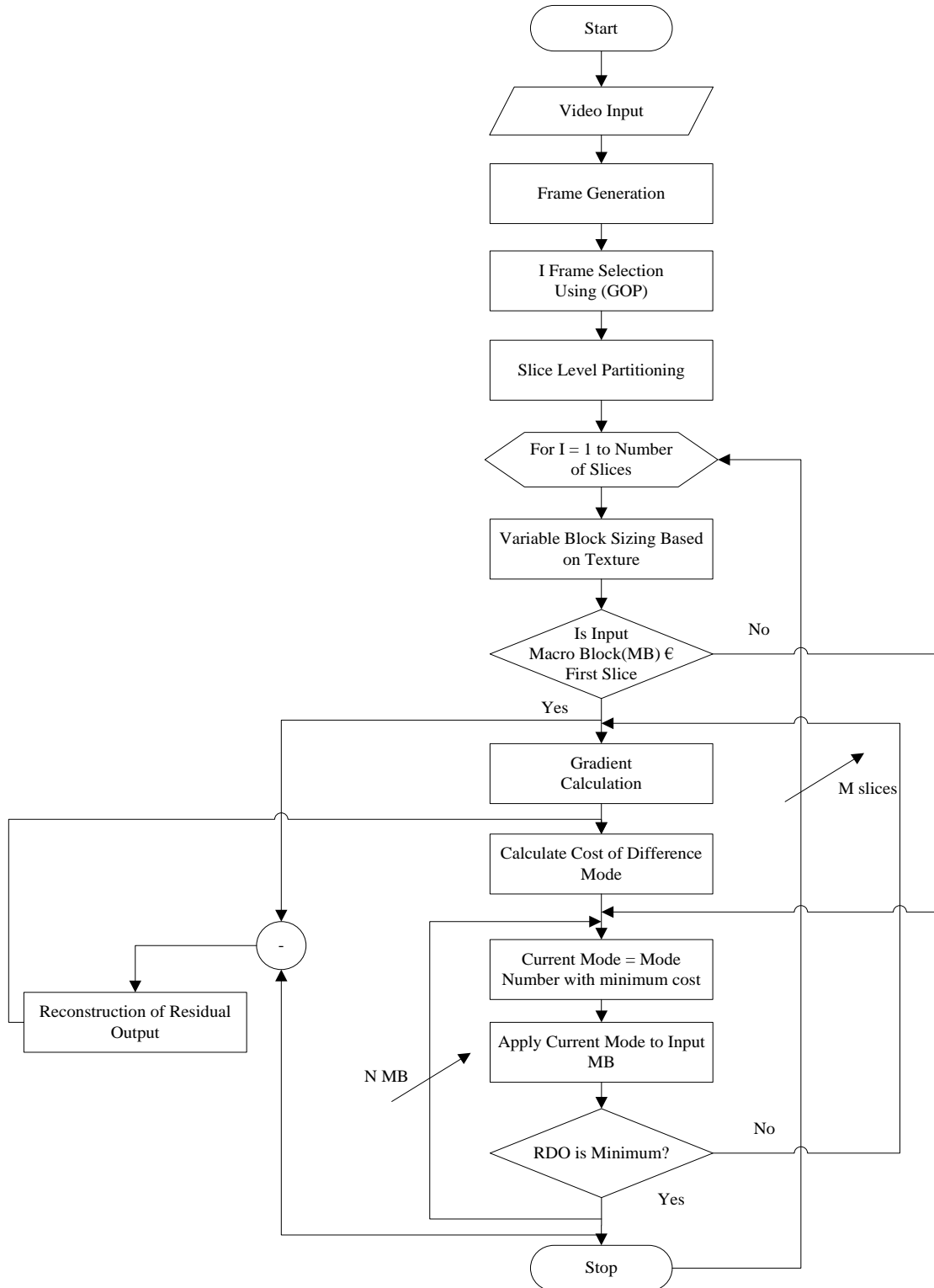


Figure 1. Block diagram of the proposed systems

An Index representing the average of the n samples in the set isy_{av} . In order to reduce the influence of inherent variation and noise terms on the sum of adjacent pixels S_{xy} , the sum of n sums of adjacent pixels is calculated by:

$$\sum_{j=0}^{n-1} S_x v(x_0, y_0 + j) \cong S_x g(x_0, y_{av} + j) \cdot \sum_{j=0}^{n-1} v(x_0, y_0 + j) \quad (6)$$

By dividing Eq (5) by Eq (6), following partial derivative can be derived:

$$\frac{\partial \ln g}{\partial x}(x_0, y_{av}) = 2 \frac{\Delta_x v(x_0, y_0 + j)}{S_x v(x_0, y_0 + j)} \cong 2 \frac{\Delta_x g(x_0, y_{av} + j)}{S_x g(x_0, y_{av} + j)} \quad (7)$$

Derivatives in the x-direction along y_{av} are calculated via Eq(7) to calculate the gradient line $g(x)$ along y_{av} . To calculate the gradient line $g(x)$, use a step-wise integration. The first point $g(x_1)$ is set to unity, and each subsequent point $g(x+1)$ is calculated from the partial derivative and the previous point $g(x)$. Complete gradient line along is produced by step-wise integration. After the gradient calculation bit cost for all the selected modes are calculated using sum of absolute difference. And the mode with minimum cost is applied to the input MB and rate distortion optimization calculation is done to that MB. If the RDO output gives minimum value then the current mode is selected as best mode and applied to the rest of the MB in the input slice. RDO calculation is done every time to every mode applied MB. If RDO calculated value gives maximum no then the gradient calculation is done again to get the best mode. Output MB is feed back to SAD calculation block every time for bit cost calculation. This process is continued for every slice in every frame until a compressed video output is obtained.

2.4 Rate Distortion Optimization

For choosing the coding mode of a macro block based on the rate and distortion cost, rate distortion optimization (RDO) mode selection technique is used. The bitrate cost R and distortion cost D is combined into a single cost J and the equation is given by:

$$J = D + \lambda R \quad (8)$$

The mode that minimizes this joint cost J is selected as best mode by RDO calculation. Lagrange multiplier (λ) controls the tradeoff between Rate and distortion. Larger λ will lead to higher distortion and reduced Rate and small λ will increase rate with less distortion. Hence choosing proper value for λ is very important and also complex. Empirical approximations have been evolved which could provide an effective λ value in practical mode selection scenario. By calculating λ as a function of QP good results can be obtained.

$$\lambda = \frac{0.852(QP - 12)}{3} \quad (9)$$

3. RESULTS

Figure 2 depicts the original frame image. The Intra coding block size is dependent more on the smoothness of the block. For an MB with detailed information the smaller block is well suited. For a smooth one bigger block is well suited.

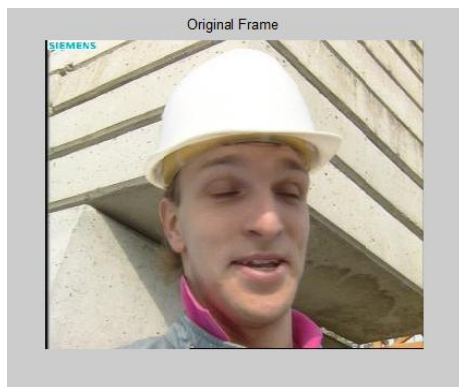


Figure 2: Original frame image

Figure 3 shows example frame from the foreman sequence. The different size of white boxes overlay on the images represent the different intra coding block size that are determined by RDO in H.264/AVC. The correlation between the smoothness and the intra coded block size can be seen.

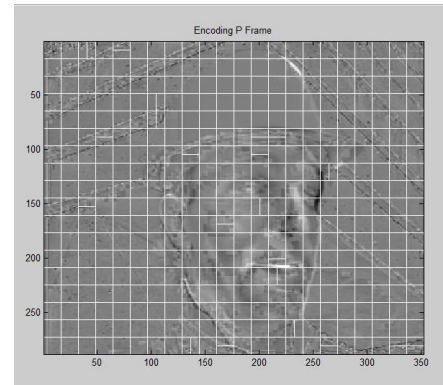


Figure 3: Example frame with intra coding blocks.

Figure 4 depicts the RD curve obtained using different QP values say, 10, 15, 20, 25 and 30 as in Table 1.

Table 1: Parameter comparison table for Existing and Expected

Methods	Measurements		
	QP	PSNR	BR
Existing	10	52.57	128.76
	15	48.74	92.23
	20	44.54	63.08
	25	40.65	40.54
	30	37.17	22.23
Expected	10	52.70	128.76
	15	48.87	92.23
	20	44.77	63.08
	25	40.71	40.54
	30	37.25	22.23

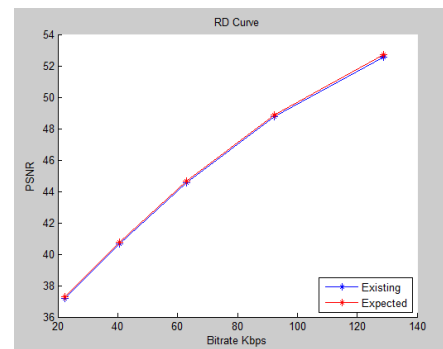


Figure 4: RD comparison curve of existing and expected outcome

Elapsed time is the total time taken by the encoder to complete all the frames. Total elapsed time taken using existing method is 387.57seconds. Proposed method is expected to take elapsed time of 370 seconds. Figure 5 depicts the elapsed time comparison graph between existing and expected methods.

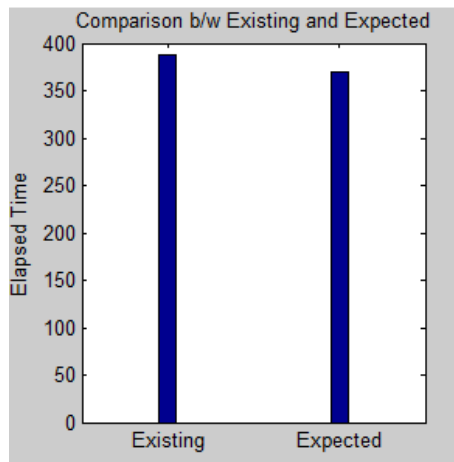


Figure 5: Elapsed time comparison graph

4. CONCLUSION

The proposed method uses an H.264/AVC video coding standard to compress the video. Fast intra prediction algorithm with gradient method is proposed in this paper. In this work number of mode combinations is expected to be reduced in every MB which takes part in RDO calculation. Due to which the computational complexity will be reduced while maintaining the bit rate and PSNR. Future research topics include the fast mode decision for the baseline profile and improvement of the RD model

5. REFERENCES

- [1] Changsung Kim, "Feature-Based Intra-/Inter-coding Mode Selection For H.264/AVC", IEEE Transactions on Circuits and Systems For Video Technology, Vol. 17, No. 4, 2007.
- [2] Y N Sairam, Nan Ma and Neelu Sinha, "A Novel Partial Prediction Algorithm for Fast 4x4 Intra Prediction Mode Decision in H.264/AVC", Data Compression Conference, Pp.232-241, 2008.
- [3] A.Elyousfi, A. Tamtaoui and E. Bouyakhf, "Fast Intra Prediction Algorithm for H.264/AVC Based On Quadratic and Gradient Model", International Journal of Electrical and Electronic Engineering, 2010.
- [4] Xiulian Peng, Jizheng Xu and Feng Wu, "Directional Filtering Transform For Image/Intra-Frame Compression", IEEE Transactions on Image Processing, Vol.19, No. 11, November 2010.
- [5] Mohammed Golam Sarwer and Q. M. Jonathan Wu, "Improved Intra Prediction of H.264/AVC", www.Intechopen.Com, 2011.
- [6] Hyoung-Mee Park, Ri-Zhu Jin and Joo-Hee Moon, "Fast Mode Selection Algorithm For Intra 4x4 Block of H.264/AVC", Pp. 645 - 649, 2011.
- [7] Qi Cai, Li Song, Guichun and Nam Ling, "Lossy and Lossless Intra Coding Performance Evaluation: HEVC, H.264/AVC, JPEG 2000 and 2000 and JPEG LS", Signal & Information Processing Association Annual Summit and Conference (APSIPA ASC), Pp. 1 - 9, 2012.
- [8] Bo Shen, Kuo-Hsiang Cheng, Yun Liu and Ying-Hong Wangm, "An Efficient Intra Prediction Algorithm for H.264/AVC High Profile", Journal of Computers Vol.25, No. 1, 2013.
- [9] Xiaohong Zhang, "Fast Mode Decision And Encryption Policy In H.264/AVC Frame-Skipping Transcoding", Journal of Computers, Vol. 9, No. 5, May 2014.
- [10] Changsung Kim, "Feature-Based Intra-/Inter-coding Mode Selection For H.264/AVC", IEEE Transactions on Circuits and Systems For Video Technology, Vol. 17, No. 4, 2007.
- [11] Y N Sairam, Nan Ma and Neelu Sinha, "A Novel Partial Prediction Algorithm for Fast 4x4 Intra Prediction Mode Decision in H.264/AVC", Data Compression Conference, Pp.232-241, 2008.
- [12] A.Elyousfi, A. Tamtaoui and E. Bouyakhf, "Fast Intra Prediction Algorithm for H.264/AVC Based On Quadratic and Gradient Model", International Journal of Electrical and Electronic Engineering, 2010.
- [13] Xiulian Peng, Jizheng Xu and Feng Wu, "Directional Filtering Transform For Image/Intra-Frame Compression", IEEE Transactions on Image Processing, Vol.19, No. 11, November 2010.
- [14] Mohammed Golam Sarwer and Q. M. Jonathan Wu, "Improved Intra Prediction of H.264/AVC", www.Intechopen.Com, 2011.
- [15] Hyoung-Mee Park, Ri-Zhu Jin and Joo-Hee Moon, "Fast Mode Selection Algorithm For Intra 4x4 Block of H.264/AVC", Pp. 645 - 649, 2011.
- [16] Qi Cai, Li Song, Guichun and Nam Ling, "Lossy and Lossless Intra Coding Performance Evaluation: HEVC, H.264/AVC, JPEG 2000 and 2000 and JPEG LS", Signal & Information Processing Association Annual Summit and Conference (APSIPA ASC), Pp. 1- 9, 2012.
- [17] Bo Shen, Kuo-Hsiang Cheng, Yun Liu and Ying-Hong Wangm, "An Efficient Intra Prediction Algorithm for H.264/AVC High Profile", Journal of Computers Vol.25, No. 1, 2013.
- [18] Xiaohong Zhang, "Fast Mode Decision And Encryption Policy In H.264/AVC Frame-Skipping Transcoding", Journal of Computers, Vol. 9, No. 5, May 2014.
- [19] Chitradevi S And Vijayalakshmi K, "Efficient Document Compression Using Intra Frame Prediction Technique", IJRET: International Journal of Research in Engineering and Technology, Vol .03, N0.07, 2014.