Original Article

Multi-Mode Schema for an Efficient Intra-Frame Prediction Technique for H.264

Pradeep Kumar N.S.[1]* and H.N. Suresh[2]

[1]*Department of Electronics & Communication Engg., South East Asian College of Engineering and Technology, Bangalore, India*
[2]*Dept. of Electronics & Instrumentation Engg., Bangalore Institute of Technology, Bangalore, India*

***Corresponding Email:** pradeepkumarnsvtu@gmail.com*

ABSTRACT

With an increasing adoption of H.264 in advance multimedia application, it is still important to understand the depth of this coding standard. Although by adopting H.264, it is ensured that compression of heavier multimedia files is done by eliminating spatial as well as temporal redundancies in order to conform to the standard, but it is also important that such video standard should ensure an efficient intra-frame coding mechanism. Therefore, the proposed system discusses about a technique that performs a simulation test-bed to evaluate Intra-frame prediction mechanism using H.264. The proposed system was also evaluated with multiple mode of selection of intra-frame prediction on various macroblock. A standard dataset is considered for evaluating the effectiveness of the proposed technique. Compression ratio and PSNR were chosen as the parameters to evaluate the performance.

Keywords: H.264, Intra-frame coding, Compression ratio, PSNR, Advance video coding.

INTRODUCTION

With an increasing usage of multimedia based application, a multimedia file (generally in the form of video) is being send from one to another terminal of IP based network. However, there are lots of differences between the file systems being transferred in wired network compared to wireless network. There are various extrinsic factors in wireless network due to signal attenuation, delay, latency, scattering, fading, interference, noise that results in potential degradation of data delivery process[1]. Usually reliability factor for bandwidth usage is highly heterogeneous especially in wireless network for which reason the network are usually shrouded by various problems associated with Quality-of-Services[2]. Owing to such significant problems in wireless network, various application witnesses considerable loss of data packet and sometimes, it results to different buffering time for multiple users for same video file itself. This significant issues in wireless network adversely affects the real-time application which usually is studied with respect to video and voice transmission. With the evolution of 3G and 4G network, individuals are considering the faster adoption of smartphones and tablet to make video calls. However till date making video calls in 3G network has not found much good response with respect to value for money they spend on it. Hence, there is a reason to perform investigation toward such issues considering multimedia file system as they are being increasingly used by the consumers in various mobile networks. However, as a matter of fact there is a solution for this in the existing era of telecommunication. With an advent of standard H.264 or advance video coding, performing video compression is no longer a big issue, but is still an issue in many cases[3]. One of the biggest problems with H.264 is that sometimes the amount of loss is found to be imperceptible. At present, H.264 has found its applicability in various multimedia applications like Blu-ray disc, YouTube, iTunes, etc. It is studied that H.264 provides enhanced encoding schemes and thereby enhances the compression and signal quality of the video on the other end i.e. receiver. However, H.264 is expected to be on rise for upcoming multimedia applications. However, such usage is also questionable as there are some potential advantages as well as limiting factors of H.264. It is important point of discussion as limiting factors as well as beneficial factors of H.264 usage are very much dependent on the evolution of upcoming high definition cameras in smartphones as well as in Tablets. It was seen that existing HD cameras and their manufacturer doesn't adopt the advance codec system like H.264 as compared to normal SD cameras. It is questionable as with the exponential rise of Smartphone user with HD cameras, H.264 may fall back. Another impediment factor is that usually the existing as well as upcoming HD cameras uses higher version of zoom for which reason, there is a higher level of dependency on computational resources to incorporate H.264 in upcoming HD cameras. And the prime cause will be failure to perform intra-frame prediction which is one of the most exclusive principles carried out by H.264 on HD videos. Therefore, in order to confirm the efficient intra-frame prediction, there is a need of evolving up with novel computational schemes that are equally compatible with hardware design.

Hence, the proposed paper discusses one such cost-effective and yet a simple technique that adopts H.264 for performing advance video coding by ensuring an efficient intra-frame prediction. The design of the proposed system is evaluated on multiple selection modes with respect to variational values of PSNR and Compression Ratio. Section 2 discusses about the prior research work that have also adopted similar principles for addressing compression issues followed by proposed system in Section 3. Section 4 discusses about the implementation technique followed by Result discussion on Section 5. Finally, some concluding remarks are being made in Section 6.

RELATED WORK

The video compression schemas are mainly based on method of motion compensated prediction process during these process realization of Scalability is difficult hence[4] presenting the work on explaining of MCTF methods. They experimentally result gives 3D wavelet demonstration, and also this paper outcomes the context of encoding digital cinema materials.

The use of transform quantization base coding for compensating the motion in video coding standards gives outstanding inaccuracy therefore[5] carried work on the investigating the effects of eliminating the encoding of bidirectional prediction residue in H.264/AVC. The experimental results shows at low to medium bit rates major improvement in the performance of the unique encoder. Data hiding mainly keen on a digital cover through a smallest amount of perceivable degradation throughout the processing of the image thus[6] have presented a data hiding scheme using H.264/AVC technique which exploit the IPCM encoded macro blocks through the intra prediction phase inside the arrangement for hiding the desired data. This method is known as blind data hiding the results outcomes low complexity method for data hiding.

Certain prior research work was also found to emphasize on the adoption of Turbo codes as well as advanced versions of wavelet. One of such work has been carried out by Ponchet and Iano[7]. Studies for adoption of region of interest was witnessed in the work done[8] where the authors have presented a unique encoder design for carrying out intra-frame prediction[9] have focused on much advanced version of H.264 by proposing an enhanced version of inter-layer prediction policy for accomplishing better bit-depth scalability[10] have focused on rate-distortion theory and introduced a framework for resource allocation for advanced version of H.264[11] have used entropy coding on H.264 standard[12] have presented a model for performing extensive interpolated structured for extracting the interpolated filters from the prior frames[13] have proposed a framework for prediction mechanism on 4x4 blocks in DC mode[14] have introduced an algorithm for investigating the inherent characteristics of human visual system. The authors have used video coding with rate-distortion theory[15] have proposed a set of techniques for enhancing the side information as well as intra-frame coding for multimedia files[16] have carried out investigation on H.264 protocol using Bayesian framework.

The prior literatures have witnessed various schemes and techniques to upgrade as well as to scrutinize the H.264 standard using various techniques. From this section, it can be seen that there are various research work priorly focusing on H.264 to enhancing the standard by more focusing on the computational complexity, whereas very few studies were witnessed with more emphasis on enhancing the coding efficiency. Hence, for this purpose, it is essential to understand that intra-frame prediction should be the pivotal point of investigation where the emphasis should be given to performing intra-frame prediction on multiple modes of selection. Such issues are addressed in the proposed paper.

Proposed model

The prime purpose of the proposed system is to design a multi-mode selection technique for Intra-frame prediction using H.264. The proposed system will apply the principle of H.264 that considers both forward and reconstruction path for data flow. The proposed system considers and input frame for the purpose of encoding, where each and every frame is subjected to processing in both inter and intra mode. In the due course of H.264 encoder design, the predicted macroblock is created depending on the newly generated reconstructed image. However, in the proposed system intra-frame coding is performed by formulating the samples of predicted macro-blocks in the current frame, which was priorly encoded and were subjected to reconstruction. Finally the predicted macroblock is subtracted from the existing macroblock to generate the remnant macroblocks. Fig. 1 exhibits the schematic diagram of the presented technique.

The proposed system also performs various mathematical operations like transformation, quantization, and inverse operation for both. By adopting scalar quantization technique in the proposed system, the scaling is performed. The proposed system also considers multiple directions for performing intra-prediction, where the outcome of the study was evaluated with PSNR value. The evaluation of the proposed system is performed in standard datasets. The implementation and results accomplished are discussed in next section.

Implementation

The proposed system is implemented in standard 32 bit Windows OS with 4 GB RAM considering Matlab as the programming tool. The implementation part of the proposed system is carried out considering H.264 encoder design principles that adopts the mechanism of intra-coded macroblocks. This step is considered for the purpose of minimizing larger quantity of data that are coded by input of original signal. In order to formulate this step, a macroblock is designed using matrix principle in Matlab that is based on priorly established reconstructed image blocks. Ultimately, using these attributes the remaining image between the existing blocks and the predicted block is subjected to encoding. The algorithm that was used to perform intra-prediction coding in proposed system is as follows:

Algorithm

Intra-prediction coding of image.

Input

Original image in .tiff format.

Output

Reconstructed image.

Start

1. Read Image I.
2. Extract Y, Cb, Cr
3. Perform Downsampling of Cb and Cr
4. Initialize Quantized and inverse quantized matrix
5. Allocate Buffer for Reconstructed Image Ir.
6. For mode i=1:8
6.1 Perform Transformation
6.2 Perform Quantization
6.3 Perform Inverse Quantization
6.4 Perform Reconstruction
7. Repeat Step 6 for Y and Cb components
8. Perform Interpolation for Cb and Cr components
9. Convert y, Cb, Cr to RGB
10. Calculate mean $PSNR_{y, Cb, Cr}$.
11. Evaluate Compression (C)
C=Image-Size / BitStream $Length_{y, Cb, Cr}$

End

For the purpose of accomplishing optimal quality of image as well as better compression ratio, the implementation of the proposed system is designed using H.264. Another prominent reason to adopt H.264 is its supportability of multiple heterogeneous sizes of block. According to the present standards of H.264, there are multiple types of prediction modes specific to multiple types of blocks. (See table 1.)

Usually, the analysis of the predicted image components are carried out by observing the behaviour of reconstructed pixels of priorly encoded adjacency matrix of the image pixels. The proposed system considers 8 modes of prediction as well as one DC mode in every sub-macroblocks of 4x4 viz. i) vertical, ii) horizontal, iii) diagonal down left, iv) horizontal down, v) diagonal down right, vi) vertical left, vii) horizontal up, and viii) vertical right. The evaluation is done by using weighted mean for the predicted samples extracted from each 8 modes in the proposed system. The design schema of the proposed technique used for intra-prediction modes is highlighted in Fig. 2.

The proposed system is designed over foreman database[17] where the multimedia file of CIF format is basically acting as input for the system. The proposed system has considered a particular image sequence from the foreman database, where a frame 512x512 in .tiff file with highest resolution of 96 dpi is considered for evaluation. In order to process this, a function is created that reads the specific format of the input frames and performs extraction of intra-frames. The proposed system is also subjected to 6 levels of quantization steps associated to 1, 2, 4, 8, 16, 32, and 64 bits after reading the RGB version of the frame. By doing this, quantization step is performed followed by extraction of y, Cb, as well as Cr components. The operation in this phase is also accompanied by performing downsampling of all Y, Cb, and Cr components of the frame to minimize the size of the bits of intra-frame that has been extracted. The proposed system than classifies the intra-frame into multiple sub-blocks, however, it was initiated with 4x4 size. It is then subjected to H.264 encoder which performs reconstruction of the frame block that has been processed. The processed block basically act as the reference to the consecutive sub-blocks by applying the fundamental process of reversing in reconstructed path of H.264 encoder. The selecting the mode of prediction, the steps of transformation, quantization, inverse quantization, and reconstruction of the frames are performed as a basic building block of H.264 encoder design. The next step is to perform interpolation for only Cb and Cr components of the intra-frames that has been recently processed. The outcome frame is then converted from YCbCr to RGB back and quality of the signal is checked. In this entire process the modes of prediction is implied on all the blocks, which is further followed by applying variable length coding (VLC) to further perform extensible compression. The final outcome is analyzed for PSNR to check the signal quality. Following are the steps that were used for performing choosing of 8 different modes of intra-frame coding in proposed system:

1. A foreman dataset is considered as input.
2. Perform extracting frames and consider it as new input
3. Perform decomposition of the frames to intra-frames
4. Divide intra-frames into multiple test macroblocks (4x4, 16x16, 32x32).
5. Classify each intra-frame to further smaller version of it(e.g. 16x16 macroblocks→4x4 sub-blocks)
6. Perform processing of preliminary sub-blocks without any prior reconstructed block.
7. Perform mathematical transformation, quantization, inverse quantization, reconstruction, and entropy encoding using VLC.
8. Perform reconstruction of sub-blocks using reverse methods.
9. Perform subtraction of consecutive sub-block with prior reconstructed sub-block and extract remnant blocks.
10. Repeat Step-7
11. Evaluated PSNR to measure signal quality.

RESULT ANALYSIS

For the purpose of result analysis, the proposed system has considered foreman dataset, where the results have been derived and discussed in this section. The analyses include executing the function with respect to multiple frames and check the outcome of the PSNR and compression ratio. Using Matlab tool, the proposed system is analyzed for various intra-prediction modes, so that the effective modes can be understood from the outcomes. For this purpose, we have analyzed both visual as well as numerical outcomes. The analysis is initiated by taking foreman video file as the input considering the coding of the test sequences where only specific intra-frames of 5 test frames are exhibited in Table 1. The outcome accomplished from the study is from foreman.yuv frame-sequences that is in CIF format. For better and precise analysis, we consider performing the experiments in 5 different values of quantization parameters (Q). Considering a frame rate of 25 frames per second, the observation is done for both original and reconstructed frames in the foreman video. Table 1 shows the original and reconstructed frames as outcome for 5 different frame-sequences of foreman dataset. A closer look into the outcomes shows higher extent of frame sharpness with proper retention of the original image quality. This outcome can easily be inferred that proposed system retain better image quality that showed superior functionality of compression being performed by H.264. It also exhibit better retention of image contents. Table 2 shows the visual outcome accomplished.

The experiment considers analyzing the outcomes with respect to variational values of quantization factor. Fig. 3-7 shows the visual outcomes of the proposed system considering Foreman frame as reference where the outcomes were analyzed with respect to different quantization factor (0, 10, 20, and 30). All the outcomes are basically reconstructed frames with higher sharpness in the outcome frames. The outcomes also shows that proposed system doesn't require higher end quantization, as the outcomes shows that better resolution figures can be obtained till the quantization factor of 10. Although increasing quantization factor doesn't affect the image quality significantly, but still it is not required.

Along with the numerical evaluation, the outcome of the proposed system was analyzed with respect to numerical values too. Fig. 8 and Fig. 9 shows the trend of observed PSNR and compression ratio on 8 different prediction modes. The outcome shows that mode 1 (vertical) to mode-5 (Diagonal downright have better performance compared to mode-6 (vertical left). horizontal up. However Mode for vertical right doesn't shows superior performance.

The proposed system has performed an experiments considering standard foreman dataset for the purpose of performing intra-prediction of frames using standard H.264. The framework designed performs selection of multiple modes of intra-prediction and evaluated the best prediction mode. From the outcomes discussed, it can be seen that all the modes doesn't perform superiorly. The data analysis observation shows that mode-1 (vertical), mode-2 (horizontal), mode-3 (diagonal down-left), and mode-4 (horizontal down) has better performance compared to other modes of selection. The outcomes observed in these modes are found with better PSNR and compression ratio performance.

REFERENCES

1. T. Manning, "Microwave Radio Transmission Design Guide", Artech House, Technology & Engineering - 280 pages, 2009.
2. Lagkas, Thomas D., "Wireless Network Traffic and Quality of Service Support: Trends and Standards", IGI Global, Political Science - 572 pages, 2010.
3. C. Zhu, Y. Li, "Advanced Video Communications over Wireless Networks", CRC Press, Computers - 512 pages, 2013.
4. J-R. Ohma, M. V.D. Schaarb, J.W. Wood, "Interframe wavelet coding-motion picture representation for universal scalability", Elsevier, Signal Processing: *Image Communication*, Vol. 19, pp. 877–908, 2004.
5. N.M.M. Rodrigues, E. A.B D. Silva, M. B. D. Carvalho, S.M.M.D. Faria, and V. M.M D.Silva, "On overriding H. 264/AVC B-slice predicted residue coding", In Picture Coding Symposium (PCS), Lisbon, Portugal, 2007.
6. S.K.Kapotas and A.N. Skodras, "Real time data hiding by exploiting the IPCM

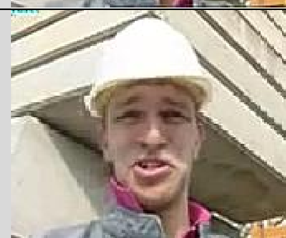
macroblocks in H. 264/AVC streams", *Journal of Real-Time Image Processing*, Vol. 4, No. 1, pp. 33-41, 2009.

7. A. F. Ponchet, Y. Iano, "Distributed Video Coding Using Turbo Codes and Second Generation Wavelets", *IEEE Latin America Transactions*, Vol. 6, No. 7, 2008.
8. F. Boulos, W. Chen, B. Parrein, and P. L. Callet, "A new H. 264/AVC error resilience model based on regions of interest", In Packet Video Workshop, 17th International, pp. 1-9, 2009.
9. L.Feng, E. Guihua, D. Qionghai, X. Hongjiang, "Bit-Depth Scalable Coding Using a Perfect Picture and Adaptive Neighboring Filter", *Tsinghua Science and Technology*, Vol. 14, No. 5, pp570-574, 2009.
10. J.Liu, Y. Cho, Z. Guo and C-C. J. Kuo, "Bit Allocation for Spatial Scalability Coding of H.264/SVC with Dependent Rate-Distortion Analysis", *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 20, No. 7, 2010.
11. S-H.Kim, J. Heo, and Y-S. Ho, "Efficient entropy coding scheme for H. 264/AVC lossless video coding", Signal Processing: *Image Communication*, Vol. 25, No. 9, pp. 687-696, 2010.
12. W. Yuping, Xiangyang Ji, Naiyao Zhang, and Qionghai Dai. "Improved Adaptive Interpolation Filter for H. 264/AVC", *Tsinghua Science & Technology*, Vol. 15, No. 2, pp. 216-220, 2010.
13. T. U. Alam, "Fast DC Mode Prediction Scheme For Intra 4x4 Block In H.264/AVC Video Coding Standard", *International Journal of Advanced Computer Science and Applications*, Vol. 3, No. 9, 2012.
14. X.Wang, H. Cui, K. Tang, "Attention-Based Adaptive Intra Refresh Method for Robust Video Coding", *Tsinghua Science and Technology*, Vol. 17, No. 1, 2012.
15. S.Rup, "Intra-Key-Frame Coding and Side Information Generation Schemes in Distributed Video Coding" Department of Computer Science and Engineering National Institute of Technology Rourkela, 2013.
16. J. Kim, J-H. Lee, B-G.Kim, J. S.Choi1, "Fast mode decision scheme using sum of the absolute difference-based Bayesian model for the H.264/AVC video standard", published in *IET Signal Processing*, 2013.
17. http://trace.eas.asu.edu/yuv/.

Table 1. Existing prediction modes

Prediction modes	Block Size	Supportability
9	4x4	Baseline
4	16x16	Main & extended profile
2	8x8	High profile

Table 2. Visualization of outcomes

Frames	Original	Reconstructed
1		
2		
3		
4		
5		

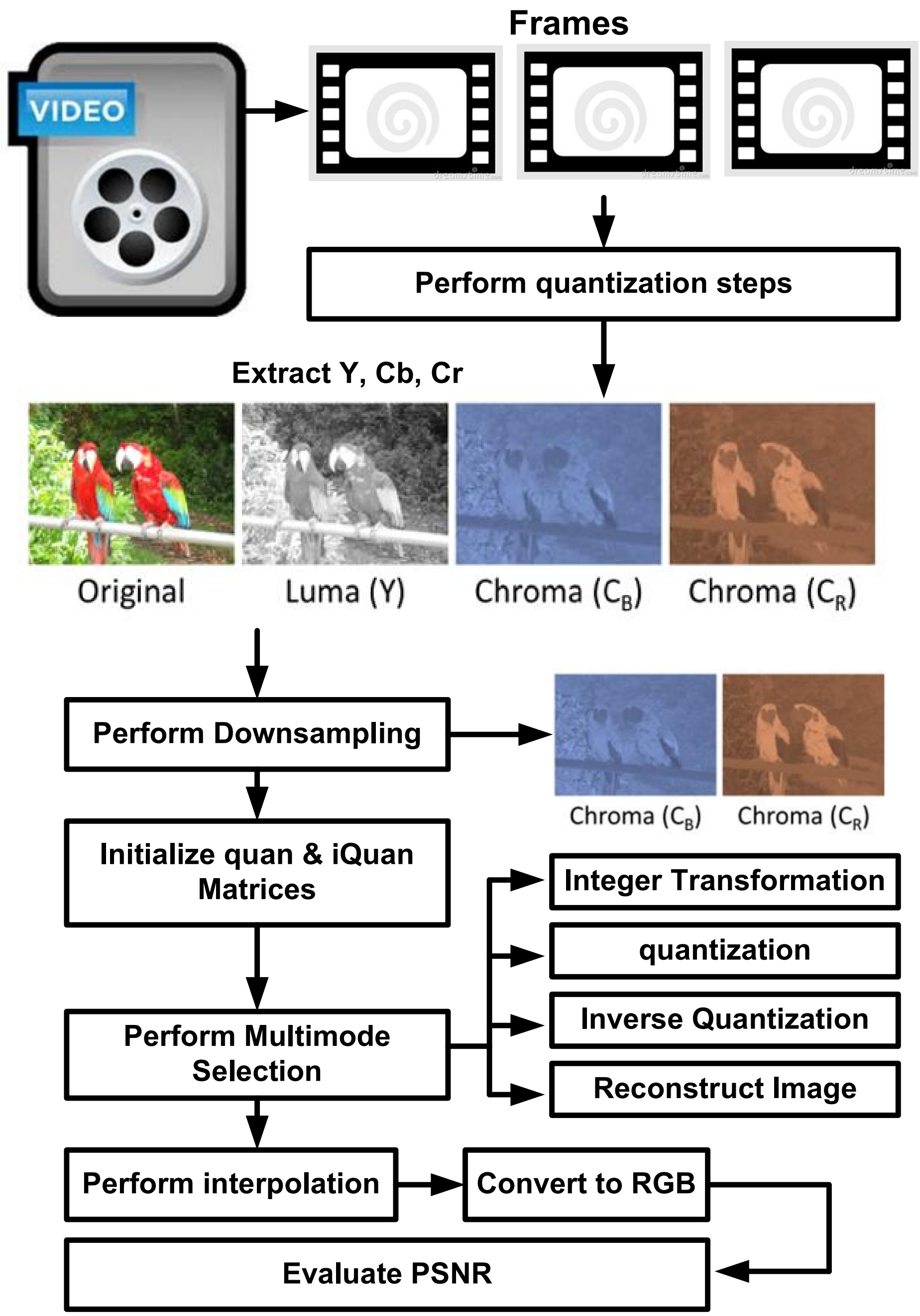


Figure 1. Proposed Schema for Intra-Frame Coding

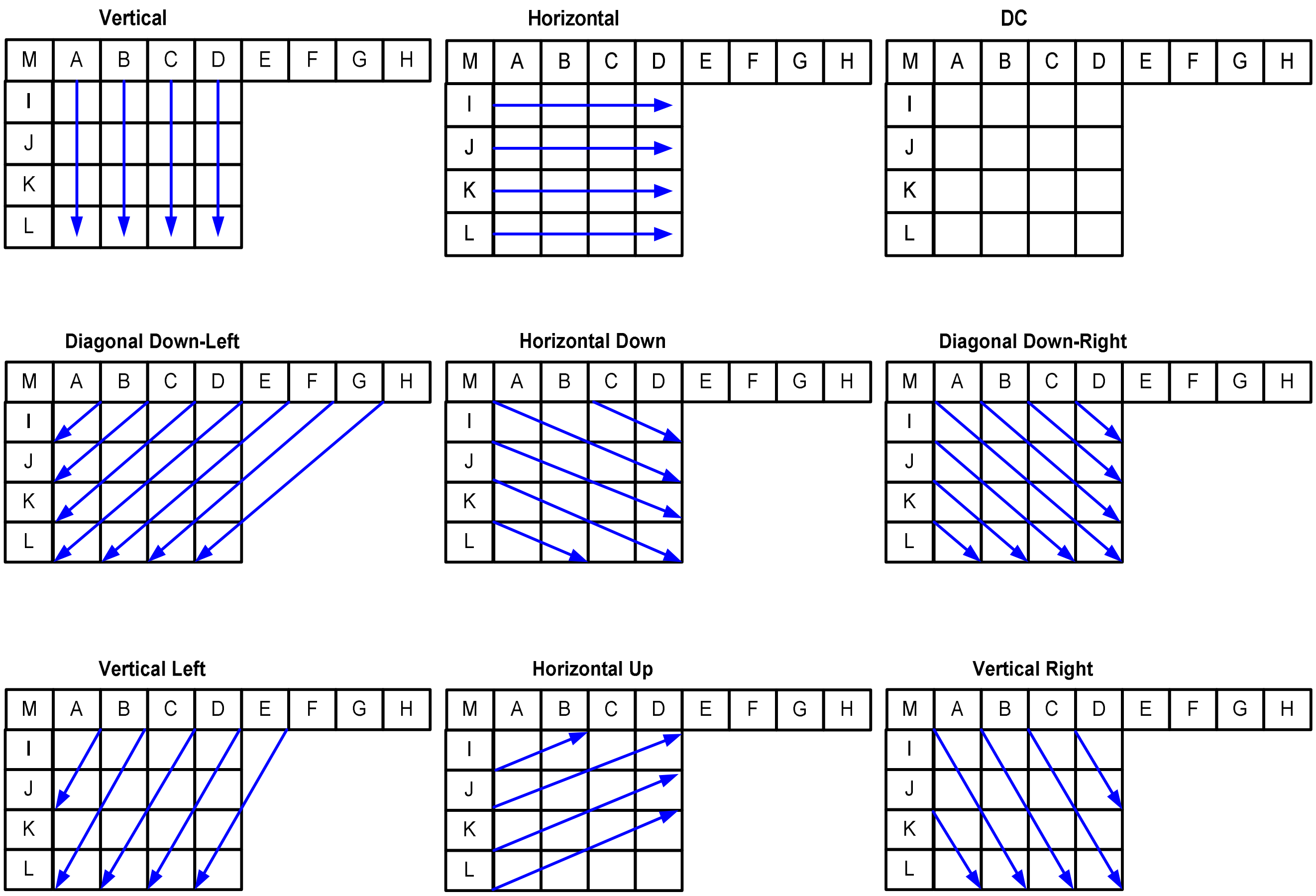


Figure 2. Formulation of intra prediction modes

Figure 3. Original frame of Foreman (input)

Figure 4. Reconstructed frame of Forman (QP=0)

Figure 5. Reconstructed frame of Foreman (Q=10)

Figure 6. Reconstructed frame of Foreman (Q=20)

Figure 7. Reconstructed frame of Foreman (Q=30)

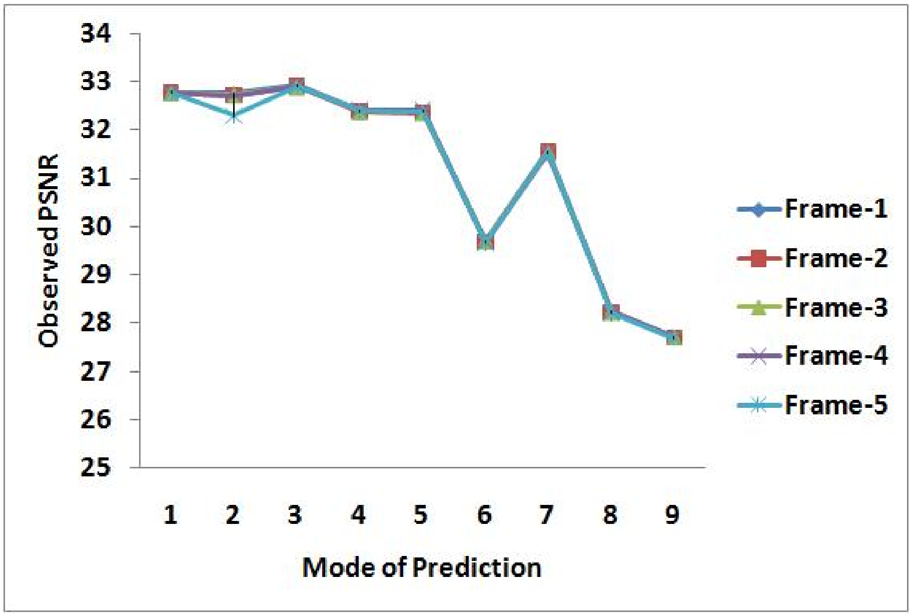


Figure 8. Analysis of Observed PSNR

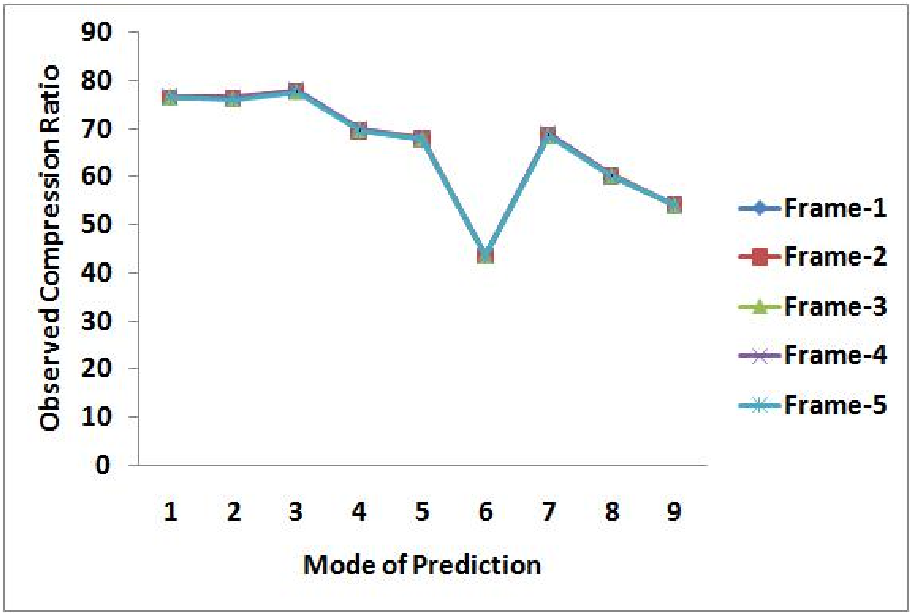


Figure 9. Analysis of Observed Compression Ratio