An Innovative Approach towards the Video Compression Methodology of the H.264 Codec: Using SPIHT Algorithms

Charu Pandey, Satish Kumar, Rajinder Tiwari

Abstract—The video signal is an integral part of multimedia which has a tremendous importance in most of the applications involving the concept of the multimedia i.e. video conferencing, video-on-demand, broadcast digital video, and high-definition television (HDTV), etc. which are expected to have a substantial user or the consumer market. It means that these applications which are based on the basic principle of the digital signals having both component of the audio and video, are using the most dominant approach of the video compression techniques so as to make the best and efficient use of the available bandwidth. The existing video coding techniques such as those used in MPEG-1 (ISO), MPEG-2 (ISO), MPEG-3 (ISO), H.261 (ITU), and H.263 (ITU) are the typical and most commonly used standards that employ the basic and most essential schemes or techniques that uses the intra frame coding or inter frame coding techniques so as to achieve the high desired compression rates. In this discussion, the authors put forward an innovative approach of the video signal compression that makes the use of the advanced H.264 algorithms which are based on the application of DCT & wavelet based video compression of the multimedia based signal. The simulation of the considered image has been carried out with a comparative analysis of the results that has been obtained with the traditional approach as well as proposed one i.e. SPIHT (Set Partitioning in Hierarchical Trees) algorithm for video compression of the signal.

Index Terms—MPEG, JPEG, Frame, Video Compression Efficiency, H.264 Algorithms, DCT, SPIHT, Video Compression, Wavelet, Block Matching, Low Bit Rate, AVC.

I. INTRODUCTION

The H.264 is the most commonly implemented traditional approach as per the present scenario of the latest generation standard for the video encoding of the signal with desired goals and efficient results. The most basic and desired goal of this approach is that it should provide good video quality at the substantially lower bit rates as compared to that obtained with the previous standards and with better error robustness i.e. better video signal quality at an unchanged baud rate. This traditional standard of the video compression of the signal has been further designed in a way so as to provide the latency of the signal as well as the better quality for a better and higher latency rate. In addition to all these improvements as compared to the previous traditional standards of the compression techniques, the improvement has been desired to be obtain without increasing the complexity of the design, in such a way so that it would not be impractical or expensive to build the applications and the systems [1]. An additional goal of this approach of the implementation of the compression of the signal has been to provide the enough flexibility so as to allow the standard to be applied to the wider range of the applications for both low and high bit rates, for low and high resolution video, and with high and low demands based on the latency rate of the data. Thus, there are a number of applications with different requirements have been identified for H.264 approach of the compression of the signal i.e.

- The entertainment video signal including the data transmission rate for the broadcast purpose, satellite communication, using cable, DVD, etc (1-10 Mbps, with high latency rate).
- The telecom services applications with <1Mbps i.e. low latency rate.
- The streaming services of the data transmission with low bit-rate and high latency of the data.

Thus, with the availability of all these features of the compression of the video signal, one can have the DVD players for the high-definition DVD formats such as HD-DVD and Blue ray supporting movies that are encoded with H.264 [2].

1) Development Stages of the H.264/AVC Algorithm:

At the end of the era of 1990s, a new group of the people working on the implementation of the various probable algorithms of the video signal compression was formed, known as the Joint Video Team (JVT) that has the most desired features of the VCEG and MPEG. The main reason behind the establishment of this joint venture was to define a standard for the upcoming next generation of the video coding of the multimedia signal. When this joint venture group has been established somewhere in May 2003, the resultant effect of this was the simultaneous launching of the same approach with high level of the recommendation from the International Telecomm Union (ITU) i.e. “ITU-T recommended for the H.264 as the highly efficient and the most reliable approach of the advanced video coding of the signal for the generic audiovisual services and thus, establishes as a standard by ISO/IEC (“ISO/IEC 14496-10 Advanced Video Coding”) [3].

During the implementation of the video compression algorithm for the multimedia signals the term “MPEG-4 part 10” has been used which generally refers to the fact that the ISO/IEC standard based on the MPEG-4 part 10, actually consists of many parts or sub modules with the current one being the MPEG-4 part 2 module. Thus, the latest and the new standard of the implementation of the algorithms was developed by the JVT and got added to the MPEG-4 as a separate part i.e. part 10 and known as “Advanced Video Coding” (AVC). Thus, based on this complete scenario of the discussion, this approach is commonly abbreviated as AVC that has been stems from the H.264 which is
An Innovative Approach towards the Video Compression Methodology of the H.264 Codec: Using SPIHT Algorithms

the resultant effect of a joint project between the ITU-T’s Video Coding Experts Group and the ISO/IEC Moving Picture Experts Group (MPEG). The ITU-T is the complete reference of the Volume 2, issue 1, January 2012 sector that coordinates the telecommunications standards on behalf of the International Telecommunication Union (ITU), with the collaborative work of the ISO which stands for International Organization for Standardization and IEC which stands for International Electro technical Commission, that is an overseas standards for all electrical, electronic and related technologies [4].

Fig1: Structure of H.264/AVC video encoder.

The above fig1 shows the brief technical structure of the H.264/AVC based encoding approach of the video signal. An alarming increase in the number of the services and the growing popularity of the high definition signal based applications of the television are creating greater and dire needs for the implementation of the higher end coding efficiency of the multimedia signals.

In addition to this approach of the video compression, there are certain other aspects of the transmission of the audiovisual signals through the possible media such as the cable modem, DSL or UMTS that offers much lower data rates than the broadcast channels and enhanced coding efficiency, which can enable the transmission of more video channels or higher quality video representations within the existing digital transmission capacities [5]. The fig2 shows the complete methodology of implementing the coding structure for the H.264/AVC algorithm for a macro block i.e. set of the multimedia data. The video signal coding for the telecommunication based applications has been evolved through the development of the ITU-T H.261, H.262 (MPEG-2), and H.263 video coding standards (and later enhancements of H.263 known as H.263+ and H.263++) that has been diversified from ISDN and T1/E1 service to embrace PSTN, mobile wireless networks, and LAN/Internet network delivery[6].

2)  Technical features of H.264:

The technical and the most desired features which are required for the design & implementation of the H.264 based algorithm of encoding the multimedia based video signals, which in turn ensures or enables the satisfactory operation of the various applications are discussed below [8]:

- The variable block-size motion compensation with small block sizes.
- The quarter-sample-accurate motion compensation
- Context-adaptive entropy
- The data Partitioning
- The decoupling of referencing order from display order
- The weighted prediction
- The decoupling of picture representation methods from picture referencing capability
- The improved “skipped” and “direct” motion inference
- The directional spatial prediction for intra coding
- The short word-length transform
- The flexible slice size

The SPIHT Algorithm:

Fig 4: SPIHT Video Coding System

The SPIHT video coding technique is shown in above fig 4 i.e. the set partitioning in hierarchical trees (SPIHT) and is popularly, known as an image compression algorithm approach that exploits the inherent similarities across the sub bands in the wavelet decomposition of an image with effective results. It is the most commonly used one in the multi-resolution pyramid format of the data which is available after the sub-band/wavelet transformation has been performed.

The inbuilt embedded coding property of the data using the SPIHT algorithm which allows the exact bit rate control without any penalty in the performance i.e.
MSE distortion control rate. The SPIHT algorithm basically, encodes the individual bits of image wavelet transform coefficients, thereby, following a bit-plane sequence. Hence, this algorithm based encoding of the signal is highly capable of recovering the image perfectly by coding all bits of transform [9 - 12].

II. WORKING STEPS OF THE SPIHT ALGORITHM

**STEP 1: Transformation of image into wavelet:**
Initially the image is transformed into wavelet domain using Wavelet Decomposition technique[16].

**STEP 2: Obtaining Wavelet Coefficients:**
The Wavelet coefficients are obtained from Wavelet Decomposition, used as an input to the SPIHT compression.

**STEP 3: Coding of Multi wavelet Coefficients Using Modified SPIHT Algorithm:**
Finally Compressed Image is obtained. Considering that the actual implementation of the SPIHT algorithm is not known, the project is an attempt to make the act clearer. For evaluating the effectiveness of multi wavelet transform for coding images or videos at low bit rates, an effective quantization and embedded coding of the coefficients has been realized. Some of the embedded coding schemes are embedded image coding using zero trees of Wavelet coefficients (EZW), and SPIHT. The SPIHT coder is a powerful image compression algorithm that produces an embedded bit stream from which best reconstructed images can be obtained at the various bit rates.

This algorithm thus improves the perceptual quality of image at all bit rates. The modified SPIHT algorithm for multi wavelet coefficient differs from the ordinary SPIHT algorithm by the way in which the subsets are partitioned and significant information is conveyed. A tree structure, called spatial orientation tree, defines spatial relationship on hierarchical pyramid. Figure 9 shows how the spatial orientation tree is defined for multi wavelet coefficients. The following sets of coordinates are used to present the new coding method [14 - 17]:

- O (i, j): Set of coordinates of all the offspring of nodes (i, j)
- D (i, j): set of coordinates of all the descendants of node (i, j)
- H (i, j): set of coordinates of all the spatial orientation tree roots (nodes in highest pyramid level)
- L (i, j): D (i, j) – O (i, j) (all the descendents except the offspring).

III. IMPLEMENTATION OF THE MODIFIED SPIHT CODING ALGORITHM

All the steps of the modified SPIHT coding algorithm are same except the initialization process. The modified SPIHT algorithm can be summarized as follows:

Fig8: Scanning and quantization order of sub images of Multi wavelet decomposition
Since the order in which the subsets are tested for the significance is important, in a practical implementation, the significance information is then stored in the ordered list called (a) List of Insignificant Sets (LIS), (b) List of Insignificant Pixels (LIP), (c) List of
Significant Pixels (LSP). In all lists each entry is identified by the coordinate \((i, j)\), which in the LIP and LSP represents individual pixels, and in the LIS represents either the set \(D(i, j)\) or \(L(i, j)\). To differentiate between them, we say that a LIS entry is of type A. If it represents \(D(i, j)\), and of type B if it represents \(L(i, j)\) [18].

Steps of the implementation of the algorithms:

1) Initialization of the process as given by the below equation 1 i.e.
\[
n = \left\lceil \log_2 \left( \max(i, j) \right) \right\rceil \left\lceil |C_i, j| \right\rceil
\]
where \(i, j \in C\) is called wavelet coefficient.

2) Add the coordinates of \(H\) in which scanning order of sub-bands as shown in figure 5, to the LIP and only those with descendents also to the LIS, as type A entries in the same order. For example the order of scanning of sub-bands is \(\text{‘L1L1’, ‘L2L2’, ‘H1L1’, ‘H2L1’, ‘H1L2’, ‘H2L2’, ‘L1H1’, ‘L2H1’, ‘L1H2’, ‘L2H2’, ‘H1H1’, ‘H2H1’, ‘H1H2’ and ‘H2H2’}\).

3) Sorting Pass:
For each entry \((i, j)\) in the LIP do:
   a) Output \(\text{Sn}(i, j)\)
   b) If \(\text{Sn}(i, j) = 1\), then move \((i, j)\) to the LSP and output the sign of \(C_{i,j}\)
   c) If \(\text{Sn}(i, j) = 0\), then add \((i, j)\) to the end of LIS as entry of type B, and go to step 4 otherwise remove entry from the LIS.

4) Refinement Pass:
For each entry \((i, j)\) in the LSP except those included in the last sorting pass (i.e. with the same \(n\)), output the \(n\)th most significant bit of \(C_{i,j}\).

5) Quantization Step Update:
Decrement the value of \(n\) by 1 and go to sorting pass if \(n\) is not less than 0.

**IV. PERFORMANCE PARAMETERS OF THE SPIHT ALGORITHMS**

**i) Mean Squared Error (MSE) and Peak Signal to Noise Ratio (PSNR).**

The two of the error notations used to compare various image compression techniques are the Mean Square Error (MSE) and the Peak Signal to Noise Ratio (PSNR) [18]. The MSE is known as the cumulative squared error between compressed and the original image, whereas the PSNR is a measure of the peak error.

The phrase peak signal-to-noise ratio is often abbreviated as PSNR, an engineering term for the ratio between maximum possible power of the signal and the power of corrupting noise that affects fidelity of its representation. Because many of the signals have a very wide dynamic range, PSNR is usually expressed in terms of the logarithmic decibel scale. PSNR is most commonly used as a measure of quality of reconstruction of lossy compression codecs [19]. When comparing the compression codecs it is used as an approximation to human perception of the reconstruction quality, therefore in some cases one reconstruction may appear to be closer to that of the original than another, even though it has a lower PSNR (a higher PSNR would usually indicate that the reconstruction is of higher quality). One has to be extremely careful with the range of validity of this metric; it is only conclusively valid when it is used to compare results from the same codec and same content [20]. It is mainly defined via the mean squared error (MSE) for two \(m \times n\) images \(I\) and \(K\) where one of the images is considered a noisy approximation of the other and is defined as [19]:

\[
MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \left( I(i, j) - K(i, j) \right)^2
\]

(2) The PSNR is defined as:

\[
PSNR = 10 \log_{10} \left( \frac{\text{MAX}^2}{MSE} \right)
\]

(3) Here, \(\text{MAX}\) is the maximum possible pixel value of image. When the pixels are represented by \(B\) bits per sample, then value is 255. Generally, when samples are represented by linear PCM with \(B\) bits per sample, \(\text{MAX} = 2^{B-1}\).

For color images with three RGB values per pixel, definition of PSNR will be same except the MSE, which is the sum of over all squared value differences divided by the image size and by three. Alternately, for color images the image is then converted to a different color space and PSNR is reported against each channel of that color space.

**ii) Compression Ratio (CR) and Bit-Per-Pixel (BPP).**

A measure of attained compression is given by the Compression Ratio (CR) and the Bit-Per-Pixel (BPP) ratio. CR and BPP represents equivalent information. CR indicates that compressed image is stored using the CR % of the initial
storage size while BPP is known as the number of bits used to store one pixel of the image.

For a greyscale image, the initial BPP is 8[21]. For a true color image the initial BPP is 24, because 8 bits are being used to encode each of three colors [20 - 21].

V. SIMULATION RESULTS & DISCUSSIONS:

Fig10: Frame no.1 for max loop 12
CR=1.3983, BPP =0.3356 CR= 2.6585, BPP = 0.6380
MSE=29.7642, PSNR=33.394 dB

Fig11: Frame no.2 for maxloop 13
CR= 2.6585, BPP = 0.6380
MSE=13.1584, PSNR = 36.939 dB

Fig.10 Frame no.1 compressed by using SPIHT algorithm
Hence it is clear that the SPIHT algorithm is highly efficient. The highest degree of flexibility in H.264’s SPIHT algorithm block-based motion compensation pays off in crowded surveillance scenes where the quality is needed to be maintained for demanding applications. Motion compensation is the most demanding aspect of a video encoder or a company and the different ways and degrees with which it can be implemented by an H.264 encoder can have an impact on how efficiently video is being compressed.

VI. CONCLUSIONS

The final output came after the research is that SPIHT algorithm provides better video compression without effecting image quality. It offers technique that enables better compression efficiency due to more accurate prediction capabilities, as well as improved resilience to errors. It provides possibilities of better video encoders that enable higher quality video streams, higher frame rates and higher resolutions at maintained bit rates (compared with the previous standards), or, conversely, the same quality video at lower bit rates. It provides a base for further research and development in the area of wavelets and general objective of achieving high degree of quality of service can be met. With the support from various industries and applications for consumer and their professional needs, H.264 is expected to replace the other compression standards and methods that are being used these days.

VII. ACKNOWLEDGEMENT

The authors are thankful to Mr. Aseem Chauhan (Additional President, RBEF and Chancellor AUR, Jaipur), Maj. General K. K. Ohri (AVSM, Retd.) Pro-VC & Director General, Amity University, Uttar Pradesh Lucknow, Prof. S. T. H. Abidi (Director ASET, Lucknow Campus), Brig. U. K. Chopra (Director AIIT & Dy. Director ASET), Prof O. P. Singh (HOD, Electrical & Electronics) and Prof. N. Ram (Dy. Director ASET) for their motivation, kind cooperation, and suggestions.

REFERENCES

1) S.M. Akramullah, I. Ahmad, M.L. Liou, A data-parallel approach for real-time MPEG-2 video encoding, Journal of Parallel and Distributed Computing 30
An Innovative Approach towards the Video Compression Methodology of the H.264 Codec: Using SPIHT Algorithms


Ms Charu Pandey is presently pursuing M.Tech in Electronics and Communication Engineering from Amity University, Uttar Pradesh, Lucknow. She has done B.Tech degree course in Electronics & Communication Engineering from Azad Institute of Engineering and Technology, Lucknow in 2011 under the affiliation of the UPTU, Lucknow. Presently, she has focused her working area within the various aspects of the image processing using some suitable simulation software in the vicinity of implementation of some highly efficient and accurate algorithms for digital watermarking.

Sh. Satish Kumar, M.Tech, FIETE is an alumnus and gold medalist of IIT Roorkee (Erstwhile University of Roorkee). He has worked as General Manager and Chief Vigilance Officer in India’s first PSU, ITI Ltd. Bangalore. He has more than 34 years of industrial experience on state-of-the-art telecom equipment and infrastructure dealing with turn-key projects. His professional experience covers the whole spectrum of switching, transmission, access, telemetry and subscriber premises equipment. He also worked for providing design, system integration and Value Engineering solutions for various network based turn-key projects like Army Static Communication Network (ASCON), Integrated Electronic Warfare System (IEWS), Defence Communication Network (DCN), Tactical Communication System (TCS), Network For Spectrum (NFS) etc., Presently he is working as Professor, (Electronics & Communication Engineering) Amity University Lucknow campus.

Mr. Rajinder Tiwari, PhD (P), M.Tech, MIETE is a member of academic staff of Department of Electronics & Electrical Engineering (ASET), Amity University Uttar Pradesh, Lucknow, where he is serving in the capacity of Asstt. Professor in the Department of Electronics Engineering (ASET). He has done M.Tech (I&CE) and M.Sc (Electronics) from NIT, Kurukshetra and University of Jammu, respectively. Presently, he is pursuing Ph.D. (ECE) from Department of Electronics Engineering, Kumaun Engineering College, Dwarahat (Almora) under Uttarakhand Technical University. Mr. Tiwari has given his contribution to the area of Microelectronics (Modeling & Simulation of the Analog CMOS Circuits for ASP Applications), Embedded System Design, Digital System Design and Process Industries Automation and Control System Design (using Graphical Programming Language with dedicated Hardware). He has published several research papers in International/National Journals/Seminar/Conference. He is associated with several technical institutions and bodies as a life member. Before taking the assignment of Amity University, Uttar Pradesh, Lucknow, he had worked in Electronics for Societal Group, CEERI, Pilani, as a Project Scientist and a Multi – National Company as a Sr. Software Engineer (Bridge Instrumentation Division). He is also associated with the successfully implementation of the Hardware and Software for number of projects undertaken by him and in organizing number of International/National Conferences and Seminars.