An Efficient Video Codec Based on H.264/AVC Standard

Chippy James
Research Scholar, VLSI and Embedded System
St. Joseph's College of Engineering and Technology
Pala, India

Manjusha Maria Alex
Research Scholar, VLSI and Embedded System
St. Joseph's College of Engineering and Technology
Pala, India

***Abstract*— Various Video compression methods are currently being used in different broadcasting technique. As technology develops demands of end users also increases. The quality of received video depends upon the compression technique that is employed at the transmitter. In this paper we are discussing about H.264 standard. It is also known as Advanced Video coding (AVC).**

Keywords- video compression, broadcasting technique, AVC or H.264, MPEG

I. INTRODUCTION

Video compression technique has been subject to technical discussion for many years. The advent of digital TV and DVD-Video has revolutionized broadcast television and home entertainment. These applications and many more were possible by the standardization of video compression technology. AVC video compression is used in many current and emerging products. It is at the heart of digital television set-top boxes, DSS, HDTV decoders, DVD players, video conferencing, internet video and other applications. These applications benefit from video compression in the fact that they may require less storage space for archived video information, less bandwidth for the transmission of the video information from one point to another or a combination of both. Besides the fact that it works well in a wide variety of applications, a large part of its popularity that it is defined in two finalized international standards, with a third standard currently in the definition process.

H.261 [1] compression technique was adopted as an international standard by the CCIT (now ITU-T) in 1990. It was designed for use in video-conferencing applications at bit rates which are integer multiples of 64kbps. It uses the 8x8 DCT blocks. At the same time, the ISO Moving Picture Experts Group (MPEG) released its first standard, which was designed to provide video (and audio) compression for storage and playback at rates of 1.0 to 1.5 Mbps. It also uses 8x8 DCT blocks. MPEG-1 [2] uses more advanced motion compensation technique including bi-directional prediction.

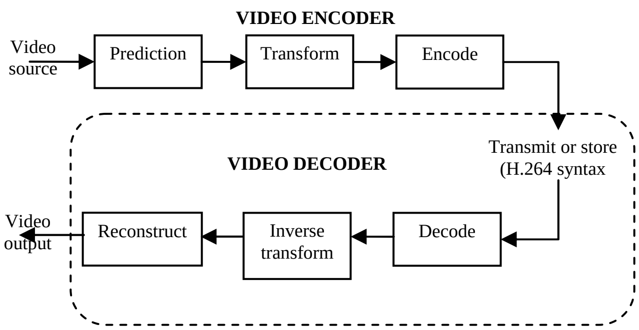
MPEG-2 [3] was developed as the standard for digital television and is currently in widespread use. It supports bit rates in the range of 2 to 10Mbps. The MPEG committee commenced work on a alternative standard for HDTV, which was due to be called MPEG-3. However this was abandoned when it became clear that MPEG-2offered reasonable support for higher resolutions. Following advances in video coding, the ITU-T released H.263 [4] as a standard for use in video telephony in1995. It was targeted at low bit rates (64kbps and under), but is suitable for higher rates as well. It also include special feature such as an advanced prediction mode.

MPEG-4 [5] standard was developed with the goal of being more than just an incremental improvement on the previous two standards. It supports a wide range of bit-rates, but is mainly focused on low bit-rate video. A fundamental concept in MPEG-4 is the idea of object-based scanning. This allows a scene to be described in terms of foreground and background objects, which may be coded independently. The ITU-T and ISO/IEC established a Joint Video Team to develop a new video compression standard using a back to basics approach. In 2003 they proposed the H.264 [6] standard, which has also been incorporated into MPEG-4 under the name Advanced Video Coding (AVC).

II. OVERVIEW OF H.264/AVC

Advanced video coding is a block-oriented motion-compensation-based codec standard developed by the ITU-T Video Coding Experts Group(VCEG) together with the ISO/IEC Moving Picture Experts Group (MPEG). It was the product of a partnership effort known as the Joint Video Team (JVT). The ITU-T H.264 standard and the ISO/IEC MPEG-4 AVC standard (formally, ISO/IEC 14496-10 - MPEG-4 Part 10, Advanced Video Coding)are jointly maintained so that they have identical technical content. H.264 is used in such applications as Blu-ray Disc, videos from YouTube and the iTunes Store, DVB broadcast, satellite television service, cable television services, and real-time videoconferencing.

H.264 standard does not explicitly define a CODEC. Rather, the standard defines a format for compressed video and a method for decoding this syntax to produce a displayable video sequence. The standard document does not actually specify how to encode digital video- this is left to the manufacturer of a video encoder – but in practice the encoder is likely to mirror the steps of the decoding process. Figure 1 shows the encoding and decoding processes and highlights the parts that are covered by the H.264 standard. An H.264 video encoder carries out prediction, transform and encoding processes to produce a compressed H.264 bit stream. An H.264 video decoder carries out the complementary processes of decoding, inverse transform and reconstruction to produce a decoded video sequence.



The encoder processes a frame of video in units of a Macro block (16x16 displayed pixels). It forms a prediction of the macro block based on previously-coded data, either from the current frame (intra prediction) or from other frames that have already been coded and transmitted (inter prediction). The encoder subtracts the prediction from the current macro block to form a residual. A block of residual samples is transformed using a 4x4 or 8x8 integer transform, an approximate form of the Discrete Cosine Transform (DCT). The transform outputs a set of coefficients, each of which is a weighting value for a standard basis pattern. When combined, the weighted basis patterns re-create the block of residual samples. The output of the transform, a block of transform coefficients, is quantized, i.e. each coefficient is divided by an integer value. Quantization reduces the precision of the transform coefficients according to a quantization parameter (QP). The video coding process produces a number of values that must be encoded to form the compressed bit stream. These values include quantized transform coefficients, information to enable the decoder to re-create the prediction, information about the structure of the compressed data and the compression tools used during encoding.

A video decoder receives the compressed H.264 bit stream, decodes each of the syntax elements and extracts the information (quantized transform coefficients, prediction information, etc). This information is then used to reverse the coding process and recreate a sequence of video images. The quantized transform coefficients are re-scaled. Each coefficient is multiplied by an integer value to restore its original scale. An inverse transform combines the standard basis patterns, weighted by the re-scaled coefficients, to re-create each block of residual data. These blocks are combined together to form a residual macro block. For each macro block, the decoder forms an identical prediction to the one created by the encoder. The decoder adds the prediction to the decoded residual to reconstruct a decoded macro block which can then be displayed as part of a video frame.

III. PROPOSED SYSTEM

The need for the compression of digital video remains undiminished despite substantial increases in bandwidth and storage capabilities over the past few years. Standards have evolved significantly during the last decade, but the basic principles of video coding remain the same. Most codecs combine two fundamental techniques: transform coding (used to compress both original and residual frames) and motion-based prediction. The proposed system has focused on the use of content-based methods of motion estimation and compensation.

In the proposed system Variable Size Block Matching (VSBM) method has been incorporated. Variable-size block matching (VSBM) allows large blocks in areas of complex motion to be split into smaller blocks (quad-tree partitioning is normally used). The use of variable-size blocks allows for greater adaptivity to local scene content. Variable Size Block Matching (VSBM) [7] provides a way of obtaining more accurate motion vectors in the case where a block straddles one or more motion boundaries.

The encoder accepts input files in Y CbCr 4:2:0 input format, with each sample represented using 8-bit precision. There is one file per frame, with the filename indicating the index of the frame within the sequence. The encoder requires the following input parameters from the user:

- The name of the Y CbCr video file to be compressed;
- The number of rows and columns of the luma component (Both of these are required to be integer multiples of 16 and less than 4096);
- The indices of the frames in the sequence at which to start and stop coding;
- The search radius used during motion estimation (This should have a value less than 256.);
- The choice of block partitioning method (Four options are available, namely: regular 16×16 blocks, regular 8×8 blocks, H.264/AVC variable size blocks, and the Binary Partition Tree block structure.);
- A quantisation parameter, QP, used in the coding of I frames, and indirectly affecting all other frames. (This allows for control over the quality and bit-rate of the compressed sequence.)

The encoder produces a compressed output file, which is used as the input to the decoder. This coded file consists of arithmetic-coded data, preceded by a seven-byte header. The header stores the video format specifications (such as resolution) and other options required by the decoder (such as the block partition choice and the value of QP). The decoder only requires one user-specified input, and that is the name of the compressed file to be decoded. All other necessary information can be obtained from the seven-byte header.

The components of the codec use Lagrangian optimisation in order to ensure that all parts of the system operate at the same trade-off gradient between rate and distortion. This is done by supplying a value of to each of the distortion. This is done by supplying a value of to each of the building blocks of the system.

Intra coding is used to code frames without reference to any other frames. The proposed video codec uses an intra frame at the start of each 15-frame Group of Pictures (GOP). The main advantages of using regular intra frames are allowing the decoder frequent access points in the bit-stream, and also helping to prevent the propagation of errors.

Intra (I) frames are coded directly using transform coding techniques, without reference to any other frames. Predicted (P) frames are motion compensated from the preceding reference frame. (Both I and P frames can act as reference frames.) Bi-directionally predicted (B) frames are motion compensated from two frames — the immediately preceding and subsequent reference frames. The proposed video codec therefore allows the option of intra blocks, which are not motion compensated from a reference frame, but instead are coded directly. Only the DC (average) values of the luma and two chroma components are coded, and no texture information is conveyed. Any significant texture is left to be coded as part of the residual image.

The process of motion compensation produces a reconstructed frame that is similar to the original. Transform coding is typically used to code the Displaced Frame Difference (DFD), which is the residual image after motion compensation. Most codecs tend to code residuals using the same transform employed for the coding of intra frames. However, the proposed video codec uses a method based on Matching Pursuit to code the residual.

Matching Pursuit is a technique that was first proposed by Mallat and Zhang [8]. It allows for a signal to be approximated using a redundant dictionary of functions. The matching pursuit algorithm does not specify a single dictionary of functions, but it is generally recommended to use a set of basis functions that match the characteristics of the signals to be approximated. Using a dictionary of basis functions an image (natural or residual) can then be approximated as follows:

- Perform matching between the image and atoms derived from the dictionary. (In Matching Pursuit, an atom is a basis function that has been translated and multiplied by a weighting factor.)
- Find the atom that minimises the total image energy when it (the atom) is subtracted from the image. Note that atoms can be translated to any location within the image, and multiplied by an appropriate weight.
- The index of this atom's basis function within the dictionary is coded, along with its translation vector and weighting factor.
- The (translated and weighted) atom is subtracted from the image, and the process is repeated. This continues until either a target number of atoms have been coded, or the image energy reaches an acceptable threshold.

Based on a 15-frame IBBPBBP. . . structure, the codec uses H.264/AVC for intra coding and then performs variable-size block matching (VSBM) to code inter frames. Intra blocks are represented using the mean values of their luma and chroma components. For each inter block, its motion vector is coded predictively, along with the choice of reference frame (in the case of B pictures). Finally, the prediction error of each inter frame is approximated and coded using the Matching Pursuit technique.

Using the above method we can encode a video sequence efficiently. This coded file consists of arithmetic-coded data, preceded by a seven-byte header. The header stores the video format specifications (such as resolution) and other options required by the decoder (such as the block partition choice and the value of QP). The decoder only requires one user-specified input, and that is the name of the compressed file to be decoded. All other necessary information can be obtained from the seven-byte header.

IV. CONCLUSION

The emerging H.264/AVC video coding standard has been developed and standardized collaboratively by both the ITU-T VCEG and ISO/IEC MPEG organizations. H.264/AVC represents a number of advances in standard video coding technology, in terms of both coding efficiency enhancement and flexibility for effective use over a broad variety of network types and application domains.

H.264/AVC has been demonstrated to provide significant rate-distortion gains over previous standards, and it is widely accepted as the state-of-the-art in video compression. As a result, efficient implementations of the standard are increasingly being used in a range of applications.

ACKNOWLEDGMENT *(HEADING 5)*

The preferred spelling of the word "acknowledgment" in America is without an "e" after the "g". Avoid the stilted expression, "One of us (R. B. G.) thanks . . ." Instead, try "R. B. G. thanks". Put sponsor acknowledgments in the unnum-bered footnote on the first page.

REFERENCES

[1] CCITT. Video Codec for Audiovisual services at p × 64 kbit/s, CCITT Recommendation H.261, 1990.

[2] CCITT. Video Codec for Audiovisual services at p × 64 kbit/s, CCITT Recommendation H.261, 1990.

[3] ISO/IEC. MPEG-1: Information Technology - Coding of Moving Pictures and associated Audio for Digital Storage Media up to about 1.5 Mbit/s – Video, Geneva, 1993.

[4] ISO/IEC JTC1 and ITU-T. MPEG-2/H.262: Generic Coding of Moving Pictures and Associated Audio Information – Part 2: Video, November 1994.

[5] ITU-T. Video Coding for Low Bit-rate Communication, ITU-T Recommendation H.263, March 1996.

[6] ISO/IEC JTC1/SC29/WG11. MPEG-4 Video Verification Model: Version 18.0, WG11 Document N3908, Pisa, January 2001.

[7] ITU-T and ISO/IEC JTC1. Advanced Video Coding for Generic Audiovisual Services, ITU-T Recommendation H.264 ISO/IEC 14496-10 AVC, 2003.

[8] M. Chan, Y. Yu, and A. Constantinides. Variable Size Block MatchingMotion Compensation with Applications to Video Coding. IEE Pro-ceeding on Communication, Speech and Vision, 137(4):205–212, August 1990.

[9] S. Mallat and Z. Zhang. *Matching Pursuit with Time-Frequency Dictionaries*. IEEE Transactions on Signal Processing, 41(12):3397–3415,December 1993.

AUTHORS PROFILE

Chippy James receiced the B.Tech (Electronics and Communication Engineering) degree from St.Joseph's College of Engineering and Technology, Pala affilated to Mahatma Gandhi University, Kottayam in 2012, and presently she is doing M.Tech (VLSI and Embedded System) degree from St.Joseph's College of Engineering and Technology, Pala.

Manjusha Maria Alex receiced the B.Tech (Electronics and Communication Engineering) degree from St.Joseph's College of Engineering and Technology, Pala affilated to Mahatma Gandhi University, Kottayam in 2012, and presently she is doing M.Tech (VLSI and Embedded System) degree from St.Joseph's College of Engineering and Technology, Pala.