ISSN 2278-3091

International Journal of Advanced Trends in Computer Science and Engineering, Vol.2 , No.1, Pages : 52-54 (2013)
Special Issue of ICACSE 2013 - Held on 7-8 January, 2013 in Lords Institute of Engineering and Technology, Hyderabad

Video Compression Using H.264 - A Review

Shahin Shaikh[1], Manjusha Deshmukh[2],
[1]Senior lecturer, Abdul Razzak Polytechnic,Navimumbai, India, shahinathvani19@gmail.com
[2]Associate professor Saraswati College Of Engineering Navimumbai, India, manju0810@yahoo.com

Abstract : H.264 is the advanced video coding standard. Video is the sequence of images played with respect to time. The successive images are highly correlated with each other. Video compression algorithms take the advantage of this fact. Only the residual information is transmitted using the technique called as block based motion estimation and motion compensation[2]. MPEG1(Motion Picture Expert Group) ,MPEG2, MPEG4, H.261, H.263 are the ancestors of H.264. Work on the emerging "Advanced Video Coding"(AVC) standard now known as ITU-T(International Telecommunication Union) Recommendation H.264 and as ISO 14496(International Organization For Standards) (MPEG-4) Part 10 has dominated the video coding standardization community. The work has been stimulating, intense, dynamic, and all consuming for those of us most deeply involved in its design. The time has arrived to see what has been accomplished. The new H.264/AVC standard is designed to provide a technical solution appropriate for a broad range of applications, including broadcast over cable, satellite, cable modem, terrestrial. It find the applications in interactive or serial storage on optical and magnetic storage devices, DVDs(Digital Video Disk).,conventional services over Ethernet, LAN(Local Area Network), wireless and mobile network and mobile.

Key words : Entropy coding, intra estimation, loop filter, motion estimation, quantization (and inverse quantization), transform (and inverse transform).

INTRODUCTION

The upcoming H.264 AVC video compression standard promises a significant improvement over all previous video compression standards. In terms of coding efficiency, the new standard is expected to provide at least 2x compression improvement over the best previous standards and substantial perceptual quality improvements over both MPEG-2 and MPEG-4.

The standard, being jointly developed by ITU-T and ISO/IEC, will address the full range of video applications including low bit-rate wireless applications, standard-definition and high- definition broadcast television, video streaming over the Internet, deliver y of high-definition DVD content, and the highest quality video for digital cinema applications.

As can be seen "History of Video Standards", the ITU-T and ISO/IEC are responsible for all previous international video compression standards. To date, the most successful of these standards has been MPEG-2, which has gone on to achieve mass-market acceptance in areas such as DVD, digital television broadcast (over cable and satellite),and digital set-top box. The new H.264 standard represents the single largest improvement in coding efficiency and quality since the introduction of MPEG-2. Consequently, over time, it is expected that H.264/MPEG-4 AVC will displace MPEG-2 and MPEG-4 ASP in many existing applications.

Technical Over view:

As can be seen in the "H.264/MPEG-4 AVC – Over view Block Diagram", the new standard is composed of several processing stages:

- Motion Estimation and Intra Estimation
- Transform (and Inverse Transform)
- Quantization (and Inverse Quantization)
- Loop Filter
- Entropy Coding

Figure 1 shows the encoder and decoder (Codec) of H.264[3].

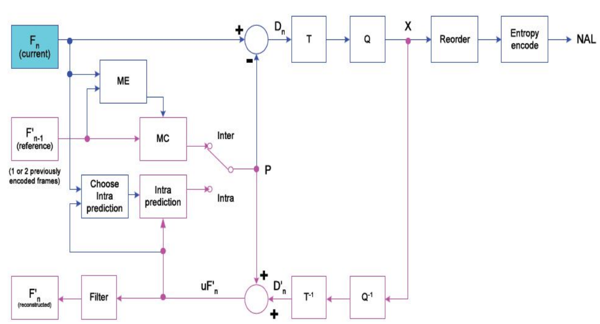
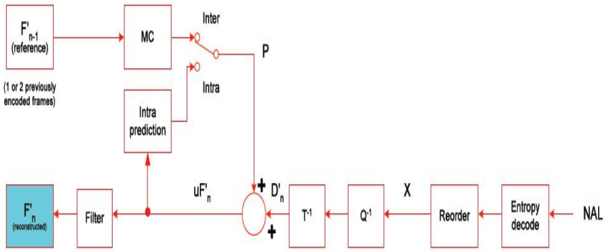


Fig . 1 H.264 CODEC

Encoder of advance video coding algorithm H.264 (Adopted from White paper: A technical introduction to H.264/AVC)

ISSN 2278-3091
International Journal of Advanced Trends in Computer Science and Engineering, Vol.2 , No.1, Pages : 52-54 (2013)
Special Issue of ICACSE 2013 - Held on 7-8 January, 2013 in Lords Institute of Engineering and Technology, Hyderabad



Decoder of advance video coding algorithm H.264 **(**Adopted from White paper: A technical introduction to H.264/AVC)

Video is composed of a stream of individual pictures that can be broken down into individual blocks of 16 pixels by 16 lines called "macroblocks". This practice simplifies the processing which needs to be done at each stage in the compression algorithm. We will explore the purpose and function of each of these processing elements in the next few Sections.

Motion estimation and Intra estimation

Motion estimation is used to identify and eliminate the temporal redundancies that exist between individual pictures. When searching for motion relative to a previous picture, the picture to be encoded is called a "P-picture". When searching both within a previous picture and a future picture, the picture to be encoded is called a "B-picture"[2].

To improve coding efficiency, the macroblock is broken down into smaller blocks that attempt to contain and isolate the motion as shown in the diagram "H.264 Motion Estimation – Superior Motion Estimation". Then, motion vectors to previous and/or future pictures are used to predict a given block. H.264 introduces smaller block sizes greater flexibility in block shapes, and greater precision in motion vectors.

Intra Estimation

In instances where motion estimation cannot be exploited, intra estimation is used to eliminate spatial redundancies. Intra estimation attempts to predict the current block by extrapolating the neighboring pixels from adjacent blocks in a defined set of different directions. The difference between the predicted block and the actual block is then coded. This approach, unique to H.264, is particularly useful in flat backgrounds where spatial redundancies often exist. An example of this is shown in "H.264 Intra Estimation".

Transform

Results from the motion estimation or intra estimation stages are transformed from the spatial domain into the frequency domain. H.264 uses a DCT-like 4x4 integer transform. In contrast, MPEG-2 and MPEG-4 employ a true DCT 8x8 transform that operates on floating-point coefficients

The smaller block size of H.264 reduces blocking and ringing artifacts. Integer coefficients eliminate rounding errors inherent with floating point coefficients and that cause drifting artifacts with MPEG-2 and MPEG-4.

Discrete Cosine Transform

The Discrete Cosine Transform (DCT) operates on **X**, a block of $N \times N$ samples (typically image samples or residual values after prediction) and creates **Y**, an $N \times N$ block of coefficients. The action of the DCT (and its inverse, the IDCT) can be described in terms of a transform matrix **A**. The forward DCT (FDCT) of an $N \times N$ sample block is given by:

$$\mathbf{Y} = \mathbf{A}\mathbf{X}\mathbf{A}^{\mathrm{T}} \qquad \text{---1}$$

and the inverse DCT (IDCT) by:

$$\mathbf{X} = \mathbf{A}^{\mathrm{T}}\mathbf{Y}\mathbf{A} \qquad \text{---2}$$

where **X** is a matrix of samples, **Y** is a matrix of coefficients and **A** is an $N \times N$ transform
matrix. The elements of **A** are:

$$A_{ij} = C_i \cos\frac{(2j+1)i\pi}{2N} \quad \text{where } C_i = \sqrt{\frac{1}{N}}\ (i=0), \quad C_i = \sqrt{\frac{2}{N}}\ (i>0)$$

Equation 1 and equation 2 may be written in summation form:

$$Y_{xy} = C_x C_y \sum_{i=0}^{N-1}\sum_{j=0}^{N-1} X_{ij} \cos\frac{(2j+1)y\pi}{2N} \cos\frac{(2i+1)x\pi}{2N}$$

$$X_{ij} = \sum_{x=0}^{N-1}\sum_{y=0}^{N-1} C_x C_y Y_{xy} \cos\frac{(2j+1)y\pi}{2N} \cos\frac{(2i+1)x\pi}{2N}$$

The output of a two-dimensional FDCT is a set of $N \times N$ coefficients representing the image block data in the DCT domain and these coefficients can be considered as 'weights' of a set of standard *basis patterns*. The basis patterns for the 4×4 DCTs is shown in Fig. 2 composed of combinations of horizontal and vertical cosine functions. Any image block may be reconstructed by combining all $N \times N$ basis patterns, with each basis multiplied by the appropriate weighting factor (coefficient)[1].

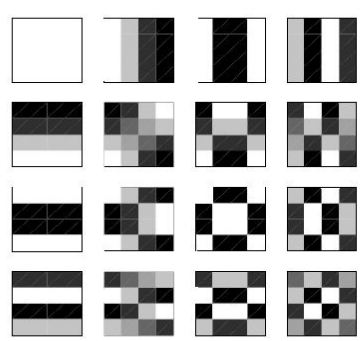


Fig. 2 : 4 × 4 DCT basis patterns

ISSN 2278-3091
International Journal of Advanced Trends in Computer Science and Engineering, Vol.2 , No.1, Pages : 52-54 (2013)
Special Issue of ICACSE 2013 - Held on 7-8 January, 2013 in Lords Institute of Engineering and Technology, Hyderabad

Quantization

The coefficients from the transform stage are quantized, which reduces the overall precision of the integer coefficients and tends to eliminate high frequency coefficients, while maintaining perceptual quality. The quantizer is also used for constant bit rate applications where it is varied to control the output bit rate.

Loop Filter

The H.264 standard defines a de-blocking filter that operates on both 16x16 macroblocks and 4x4 block boundaries. In the case of macroblocks, the filter is intended to remove artifacts that may result from adjacent macroblocks having different estimation types (e.g. motion vs. intra estimation), and/or different quantizer scale. In the case of blocks, the filter is intended to remove artifacts that may be caused by transform/quantization and from motion vector differences between adjacent blocks. The loop filter typically modifies the two pixels on either side of the macroblock boundary using a content adaptive non-linear filter[5].

Reordering

Scanning of the coefficients is called as reordering. Depending on whether these coefficients were originally motion estimated or intra estimated, a different scan pattern is selected to create the serialized stream. The scan pattern orders the coefficients from low frequency to high frequency. Then, since higher frequency quantized coefficients tend to be zero, run-length encoding is used to group trailing zeros, resulting in more efficient entropy coding[1].

Entropy Coding

The entropy coding stage maps symbols representing motion vectors, quantized coefficients, and macroblock headers into actual bits. Entropy coding improves coding efficiency by assigning a smaller number of bits to frequently used symbols and a greater number of bits to less frequently used symbols.

Variable Length Coding (VLC) and Context Adaptive Binary Arithmetic Coding (CABAC)can be used. CABAC offers superior coding efficiency over VLC by adapting to the changing probability distribution of symbols, by exploiting correlation between symbols, and by adaptively exploiting bit correlations using arithmetic coding. H.264 also supports Context Adaptive Variable Length.Coding (CAVLC) which offers superior entropy coding over VLC without the full cost of CABAC.

CONCLUSION

The new H.264offers significant bit rate and quality advantages over all previous standards. H.264 has taken a more pragmatic and focused approach to addressing the problems and need of the current and emerging multimedia applications[1]. There are the indications that H.264 may become the technical leader that will drive the next generation of digital video.

REFERENCES

[1]. Iain E.G. Richardson, "H.264 and MPEG – 4 Video Compression", *Video Coding for Next-generation Multimedia*, John Wiley & Sons LTD, 2003.

[2]. Khalid Sayood, "Introduction to Data compression", Morgan Kaufmann Publishers, 2008

[3]. Iain E.G. Richardson , *White paper: "A technical introduction" to H.264/AVC*, 2011.

[4]. Cheolhong An and Truong Q. Nguyen, "*Statistical Learning Based Intra Prediction in H.264*", IEEE, 2008.

[5]. Gulistan Raja and Muhammad Javed Mirza, "*In-loop Deblocking Filter for H.264/AVC Video*".

[6]. M. Karczewicz and R. Kurceren, "*The SP and SI Frames Design for H.264/AVC", IEEE Transactions on Circuits and Systems for Video Technology*, 2003.

[7]. D. Marpe, H. Schwarz and T. Wiegand, "*Context-Based Adaptive Binary Arithmetic Codiin the H.264/AVC Video Compression Standard*", IEEE Transactions On Circuits And Systems For Video Technology 2003.