# Issues in H.264/MPEG-2 Video Transcoding

Hari Kalva
*Computer Science and Engineering, Florida Atlantic University, Boca Raton, FL*
*hari@cse.fau.edu*

## Abstract

*The H.264 video coding standard has recently been standardized under the auspice of both the ITU and the ISO video experts groups. This new generation video standard, with its significant bandwidth savings, is expected to replace the use of MPEG-2 video compression in digital video systems. The complete migration to the new video coding algorithm will take several years given the wide scale use of MPEG-2 in the market place today. This creates an important need for transcoding technologies that transcode the widely available MPEG-2 compressed video to H.264 compressed format and vice versa. The high transcoding complexity can be reduced by reusing the information gathered in the video decoding process. However, given the significant differences between the MPEG-2 and the H.264 coding algorithms, transcoding is much more complex compared to MPEG-2 to MPEG-4 transcoding. In this paper we give a brief introduction to H.264 and discuss the issues in transcoding and complexity reduction.*

## 1. Introduction

The H.264 standard, jointly developed by the ITU-T Q.6/SG16 and the MPEG committee (ISO/IEC/SG29/WG11), is highly efficient and when compared to MPEG-2, can produce a perceptually equivalent quality video at about half the bitrate. An overview of the H.264 standard can be found in [1]. The potential of H.264 to replace MPEG-2 video in digital video systems today creates several possibilities for employing the transcoding technologies. If we consider the Cable TV industry today, we can observe that more and more cable TV systems are installing digital TV systems and digital set top boxes (STB). The move to HDTV is rather slow with one of the main reasons being the large amount of bandwidth required to support the format for relatively few users today. Using H.264 represents significant bandwidth savings for Cable TV systems creating opportunities for delivering new services including HDTV at a lower bitrate. Adopting H.264, however, will require transcoders to convert H.264 to/from MPEG-2 as necessary. One scenario is to have H.264 transcoders at the head-end of Cable TV systems to convert the input signals to H.264 or MPEG-2 depending on the end user equipment. Another area expected to see early adoption of H.264 is the hard-disk based digital video recorders that can record TV programs at a much lower bitrate. Given the relatively early stage of video capability support in mobile phones, mobile phones will be one of the first market segments to adopt H.264 video. With all these possibilities, there is a strong need for research into technologies for H.264/MPEG-2 transcoding.

In this paper we primarily discuss the research issues in the H.264/MPEG-2 video transcoding activity we have recently started. The rest of the paper is organized as follows: section 2 provides the background with a brief overview of the MPEG-2 and the H.264 video coding standards. Section 3 presents the issues in H.264/MPEG-2 transcoding. We close the paper with concluding remarks in section 4.

## 2. Background

### 2.1 Brief Overview of MPEG-2 Video

The MPEG-2 coding standard has been designed to efficiently support both interlaced and progressive video coding and produce high quality standard definition video at about 4 Mbps. The MPEG-2 video standard uses a block-based hybrid transform coding algorithm that employs transform coding of the motion-compensated prediction error. While motion compensation exploits temporal redundancies, the DCT transform exploits the spatial redundancies. The asymmetric encoder-decoder complexity allows for a simpler decoder while maintaining high quality and efficiency through a more complex encoder. In the interest of maintaining brevity, we refer you to [2] and [3] for more details on the MPEG-2 video coding standard.

### 2.2 Brief Overview of H.264 Video

The H.264 video coding standard has been developed recently through the joint work of the ITU's video coding experts group (VCEG) and ISO motion pictures experts group (MPEG). The H.264 video coding standard is flexible and offers a number of tools to support a range of applications with very low as well as very high bitrate requirements. Compared with MPEG-2 video, the H.264 video format gives perceptually equivalent video at 1/3 to

0-7803-8145-9/04/$17.00 ©2004 IEEE.

1/2 of the MPEG-2 bitrates. The bitrate gains are not a result of any single feature but a combination of a number of encoding tools. These gains come with a significant increase in encoding and decoding complexity [4]. In spite of the increased complexity, the dramatic bandwidth

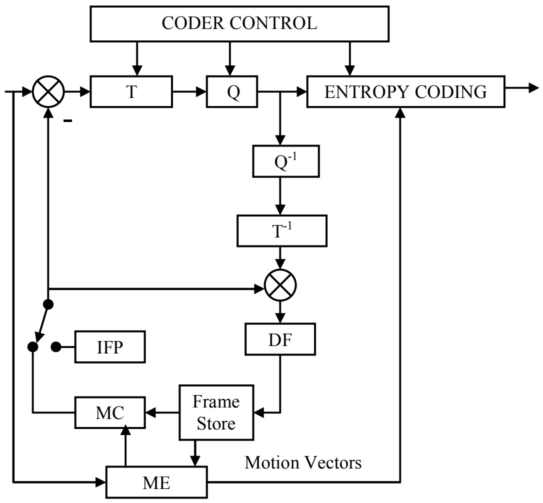


Figure 1. The H.264 Video Encoder

savings will encourage TV broadcasters to start adopting the new technology as they can now use the bandwidth savings to provide new channels or new data and interactive services. With the coding gains of H.264, full length HDTV resolution movies can now be stored on DVDs today. Further more, the fact that the same video coding format can be used for broadcast TV as well as Internet streaming will create new service possibilities and speeds up the adoption of H.264 video.

The H.264 video uses the same hybrid coding approach that is used in the other MPEG video standards: motion compensated transform coding. Figure 1 shows the block diagram depicting the main components of a H.264 video codec. The H.264 employs a hybrid coding approach similar to that of MPEG-2 but differs significantly from MPEG-2 in terms of the actual coding tools used. The main differences are: use of an integer transform (T) with energy compaction properties similar to that of the DCT instead of the DCT, an in-loop de-blocking filter (DF) to reduce block artifacts, and intra frame prediction (IFP). The coder control operation is responsible for functions such as reference frame management, coding mode selection, and managing the encoding parameter set. The H.264 standard introduces several other new coding tools that improve coding efficiency. We discuss some of these tools that are relevant to the transcoding process. A complete overview of the H.264 video standard can be found in [1].

*Intra-frame prediction:* While macro blocks (MB) of 16x16 pixels are still used, predicting a MB from the previously encoded MBs in the same picture is new in H.264. An MB can have 4x4 or 16x16 block prediction modes referred to as Intra_4x4 and Intra_16x16 respectively. There are 9 4x4 block prediction modes and 4 16x16 block prediction modes. These intra prediction modes including directional prediction and improves the prediction when directional structures are present in the picture. With the intra-frame prediction, the I-pictures can be coded more efficiently compared to MPEG-2, which does not support intra-frame prediction.

*Inter-frame prediction:* The inter-frame prediction is the traditional motion compensated prediction that is supported in earlier MPEG video coding standards. The H.264 extends this by allowing multi-frame references for prediction. When multiple reference frames are used for prediction, the encoder and the decoder synchronize the multi picture buffer using memory control operations. The motion compensation of 16x16 blocks in H.264 uses variable block sizes and motion vectors with quarter-pixel resolution. These motion compensation tools, though complex, contribute significant coding gains in H.264.

*Transform coding:* The transform coding is applied to the prediction residual as is done in MPEG-2. The transform used, however, is a 4x4 integer transform instead of the 8x8 DCT traditionally used in many video coding standards. The integer transform is designed such that the transformation involves only additions and shift operations and there is no mismatch between the forward and the inverse transform. This reduces the complexity significantly compared to the DCT. The smaller size of the transform, compared to the 8x8 DCT in MPEG-2, also reduces the block noise in the decoded video.

*Entropy coding:* The entropy coding in H.264 uses universal variable length coding (VLC) for all syntax elements except the quantized coefficients. The coefficients can be coded using either context adaptive VLC (CAVLC) or context adaptive binary arithmetic coding (CABAC). The CABAC gives the most coding gains.

## 3. H.264/MPEG-2 Transcoding

The existing transcoding research deals mainly with spatial and transform-domain reduced resolution transcoding of high bitrate video such as MPEG-2 video to a lower bitrate format, typically the MPEG-4 video format [5, 6]. The H.264 video format has just been finalized and there is no known research on H.264 transcoding to the best of our knowledge. Though the high level coding approach for MPEG-2 and H.264 encoding is the same – hybrid motion compensated

0-7803-8145-9/04/$17.00 ©2004 IEEE.

transform coding – the actual coding tools make them substantially different.

Figure 2 shows the reference architecture for MPEG-2 to H.264 transcoding. This is called the reference architecture as this represents the upper bound on the rate-distortion performance of the transcoded video. This is the simplest architecture and encompasses the complexity of the full MPEG-2 decoding and H.264 encoding. The similarities in H.264 and MPEG-2 have to be exploited to reduce the transcoding complexity. The various complexity quality tradeoffs have to be studied and various transcoding architectures have to be considered in developing realtime H.264/MPEG-2 transcoding techniques. We initially consider the following areas for complexity reduction in H.264/MPEG-2 transcoding:

- Motion Estimation and Compensation
- Intra-picture prediction
- Adaptive field/frame coding
- Transform coding

In addition to the factors discussed above, the transcoding architectures have to consider other elements of H.264 such as the slice structure, redundant frames, de-blocking filters, network adaptation layer, and entropy coding in reducing the transcoding complexity. As there are numerous combinations of coding modes possible in H.264, we will conduct research to determine the coding modes and options that lend themselves to transcoding with reduced complexity and its impact on the quality.

## 4. Conclusion

The H.264 video coding standard is flexible video coding standard and offers a number of tools to support a range of applications with very low as well as very high bitrate requirements. Compared with MPEG-2 video, the H.264 video format gives perceptually equivalent video at 1/3 to 1/2 of the MPEG-2 bitrates. In spite of its complexity, the actual bandwidth savings compared to MPEG-2 provides a big motivation to migrate to H.264 video coding. The fact that the standard is intended for use in low bitrate applications such as mobile phones, moderate bitrate application such as video conferencing, and high bitrate applications such as digital TV and digital video entertainment creates an economy of scale, resulting in cheaper hardware and providing additional motivation to migrate to H.264. Given that H.264 video has just been standardized, there is a need for significant research in transcoding technologies to enable smooth transition from MPEG-2 to H.264.

## 5. References


[1] T. Wiegand, G. Sullivan, G. Bjontegaard, and A. Luthra. "Overview of the H.264/AVC Video Coding Standard," IEEE Transactions on Circuits and Systems for Video Technology, Vol. 13, No. 7, July 2003.

[2] ISO/IEC JTC11/SC29/WG11, "Generic Coding of Moving Pictures and Associated Audio Information: Video", ISO/IEC 13818-2.

[3] B. Haskell, A. Puri, and A. Netravali, "Digital Video: an introduction to MPEG-2," Chapman and Hall, 1998.

[4] Implementation Studies Group, "Main Results of the AVC Complexity Analysis," MPEG Document N4964, ISO/IEC JTC11/SC29/WG11, July 2002.

[5] A. Vetro, T.Hata, N. Kuwahara, H. Kalva, and S. Sekiguchi, "Complexity-quality analysis of transcoding architectures for reduced spatial resolution," IEEE Transactions on Consumer Electronics, Volume: 48 Issue: 3, August 2002, Page(s): 515 -521.

[6] H. Kalva, A. Vetro, and H. Sun, "Performance Optimization of the MPEG-2 to MPEG-4 Video Transcoder," SPIE Conference on Microtechnologies for the New Millennium, VLSI Circuits and Systems, May 2003 (invited paper).


0-7803-8145-9/04/$17.00 ©2004 IEEE. 