EE5359 Project Proposal on

VIDEO COMPRESSION STANDARDS FOR HIGH DEFINITION VIDEO: A
COMPARATIVE STUDY OF H.264, DIRAC PRO AND AVS PART 2

By Sudeep Gangavati
ID: 1000717165
LIST OF ACRONYMS

AU: Access Unit
AVS: Audio Video Standard
AVS-M: Audio Video Standard for mobile
B-Frame: Bidirectionally Interpolated Frame
BBC: British Broadcasting Corporation
CAVLC: Context Adaptive Variable Length Coding
CBP: Coded Block Pattern
CIF: Common Intermediate Format
DIP: Direct Intra Prediction
DPB: Decoded Picture Buffer
EOB: End of Block
HD: High Definition
ICT: Integer Cosine Transform
IDR: Instantaneous Decoding Refresh
I-Frame: Intra Frame
ITU-T: International Telecommunication Union
JPEG: Joint Photographic Experts Group
MSE: Mean Square Error
PSNR: Peak Signal to Noise Ratio
QCIF: Quarter Common Intermediate Format
SMPTE: Society of Motion Picture and Television Engineers
SSIM: Structural Similarity Index
Video Compression Standards for High Resolution Video

Objective:
The objective of this project is to study, implement and compare video coding standards like H.264/AVC [1], Dirac pro [3] and AVS China P2 (AVS video) [8]. The analysis will be carried out and different performance metrics like MSE, PSNR, bitrate, SSIM [13] and video quality will be evaluated for high resolution videos at various bitrates.

Motivation:
With the ever increasing demand for high definition video, several different video coding standards have been developed to address the needs of HD Video coding. The project attempts to implement and evaluate the video coding standards that have been extensively used for HD video broadcasting, storage and distribution. The video coding standards that will be evaluated are H.264/AVC (Main Profile) [1], AVS China P2 (Base Profile) [8] and Dirac Pro [5]. Since Dirac Pro is intra frame coding only, the analysis will be carried out only for the intra frame coding in case of H.264/AVC [1] and AVS China P2 [8].

Introduction:
H.264/AVC

The H.264/AVC is the latest advanced video coding standard developed by ITU-T Video Coding Experts Group together with the ISO/IEC Moving Picture Experts Group [1]. It is the most widely used video coding standard [15] [28] for streaming videos, mobile/handheld applications, HDTV broadcasting etc. The H.264 standard supports three sampling patterns for luminance component (Y), red-difference chroma component (Cr) and blue-difference chroma component (Cb) [20]. The 4:4:4 sampling means that the three components (Y: Cr: Cb) have the same resolution and a sample of each component exists at every pixel position as shown in Figure 1(a).

![Fig. 1(a) 4:4:4, 4:2:2, 4:2:0 Sampling patterns](20)
An H.264 encoder converts the raw video into a compressed version and the decoder converts the compressed video back to its original format. The H.264 encoder block diagram is shown in Figure 1(b).

![Fig.1 (b) Basic H.264 encoder structure [1]](image)

The encoder performs transform, quantization, prediction and encoding to produce compressed video. The decoder shown in Figure 2 on the other hand does the inverse operations to obtain the uncompressed video.

![Fig.2 H.264 decoder structure [2] [17]](image)

**H.264 Profiles**

H.264 standard defines a set of profiles and levels to set points of conformance for different classes of applications and services. For each profile there are specific encoding tools that will be
supported by the decoders conforming to that profile. There are mainly seven profiles [2]: Baseline, Main, Extended, High, High 10, High 4:2:2, High 4:4:4. Main profile is designed for digital storage media and television broadcasting. Extended profile is designed for multimedia services over Internet. Baseline profile is aimed at real time applications such as video conferencing. High profile mainly aims at applications such as content distribution, studio editing and high resolution videos. Profiles are shown in Figure 3.

![Figure 3 H.264 profiles and levels](image)

**Video Coding Algorithm**

The encoder block diagram is shown in Figure 1 [1]. Encoder will select between intra and inter-coding for blocks of each picture. Intra coding exploits several spatial prediction modes as shown in Figure 3(a) and (b) to reduce the redundancies in the signal for a single picture. Inter coding does the inter prediction of each block of sample values from previously coded pictures. Inter coding uses motion vectors for block based inter-prediction to reduce temporal redundancy between different pictures. The deblocking filter is used to reduce the blocking artifacts at the block boundaries. The prediction residual is further transformed to remove spatial correlation in the block before it is quantized. Finally, the intra predicted modes or motion vectors are combined with the quantized transform coefficient information and encoded using arithmetic coding or entropy coding.

**Prediction in H.264**

The H.264 video coding standard employs two different prediction techniques called intra prediction and inter prediction in order to predict the current macroblock [20]. In intra prediction, the prediction for the current macroblock of image samples is created from previously coded samples in the same frame. In an intra macroblock, there are three choices for the intra macroblock size for the luma component namely 16 x 16, 8 x 8 or 4 x 4 [1][20]. A single prediction block is generated using one of a number of possible prediction modes. For a 4 x 4
macroblock there are nine modes, for 8 x 8 macroblock size available in High profile also has 9 modes but for 16 x 16 there are 4 modes [20].

During the prediction, one mode is selected and is then used to predict the values. The modes for 4 x 4 and 16 x 16 luma prediction modes are shown in Figure 3 (a) and Figure 3 (b) respectively.

Fig. 3(a) Intra prediction modes for 4 x 4 block size [20].

In inter prediction, motion estimation and motion compensation techniques are used to predict the current macroblock [20]. This process involves selecting a prediction region, generating a prediction block and subtracting this from the original block of samples to form a residual that is transformed, quantized and then encoded. The macroblock can be of different sizes as shown in Figure 3 (c).

Fig. 3 (b) Intra prediction modes for 16 x 16 block size [20].

Fig.3(c) Inter prediction macroblock sizes [20]
Deblocking Filter

A filter is applied to every decoded macroblock as shown in Figure 3(d) to reduce the blocking distortion [20]. The deblocking filter is applied after the inverse transform in the encoder before reconstructing and storing the macroblock for future predictions and in the decoder before reconstructing and displaying the macroblock. The filter smooths the block edges, improving the appearance of the decoded frames.

![Fig. 3 (d) Boundaries in a macroblock to be filtered (luma boundaries shown with solid lines and chroma boundaries shown with dotted lines)](image)

DIRAC PRO

Dirac is a video codec originally developed by BBC [3]. The main aim of Dirac video standard is to provide high-quality video compression for web streaming and HDTV applications. BBC used Dirac to transmit HDTV pictures of Beijing Olympics in 2008 [3][4]. Dirac Pro is a version of Dirac family of compression tools mainly optimized for video production and archiving applications and the focus is on high quality and low latency. Dirac Pro is intended for high quality applications with lower compression ratios [4][5].

Dirac Pro supports the following technical aspects [5]:

- Intra-frame coding only
- 10 bit 4:2:2
- No Subsampling
- Lossless or visually lossless compression
- Low latency on encode/decode
- Robust over multiple passes
- Support for multiple HD image formats and frame rates
- Low complexity for decoding
The main difference in the Dirac and Dirac Pro is the treatment in the final process in compression – the arithmetic coding. Arithmetic coding is processing intensive and introduces delay. These features are undesirable in high end production work and hence Dirac Pro omits arithmetic coding.

The encoder and decoder block diagram are shown in Figure 4 (a) and (b) respectively.

**Architecture**

Dirac can compress any size of picture from low resolution QCIF to HDTV. Dirac employs wavelet compression instead of discrete cosine transform [6] used in other codecs. Another application of wavelet transform is the JPEG 2000 compression standard for still images [7].
**Motion Estimation**

In Dirac, frames have two essential properties. Firstly, they are either predicted from other frames i.e. Inter. Secondly they can be used to predict other frames. All combinations of these properties are possible, and any inter frame can be predicted from up to two reference frames. But in Dirac Pro, only intra frame coding is used. Dirac Pro provides spatial and quality scalabilities, useful to save bandwidth during the transmission of a single bit stream to receivers with different image resolution and bandwidth requirements [6]. Dirac Pro has been adopted by SMPTE as VC-2 [29].

**AVS CHINA**

AVS is an acronym for Audio Video Standard which is a compression codec for digital audio and video developed by China [8]. AVS China was developed to replace the most used H.264/AVC standard. AVS China finds its applications in high resolution broadcast, video on wireless communications medium etc.

AVS China has been divided into various parts and thus dividing the AVS China architecture into various sub-fields. The AVS standard has been divided into 10 parts as shown in Figure 5 [16].

<table>
<thead>
<tr>
<th>Part</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>System</td>
</tr>
<tr>
<td>2</td>
<td>Video</td>
</tr>
<tr>
<td>3</td>
<td>Audio</td>
</tr>
<tr>
<td>4</td>
<td>Conformance test</td>
</tr>
<tr>
<td>5</td>
<td>Reference Software</td>
</tr>
<tr>
<td>6</td>
<td>Digital media rights management</td>
</tr>
<tr>
<td>7</td>
<td>Mobile video</td>
</tr>
<tr>
<td>8</td>
<td>Transmit AVS via IP network</td>
</tr>
<tr>
<td>9</td>
<td>AVS file format</td>
</tr>
<tr>
<td>10</td>
<td>Mobile speech and audio coding</td>
</tr>
</tbody>
</table>

Fig.5 AVS parts [16]
AVS part 1 considers the system for broadcast. AVS Part 2 considers the video part. AVS Part 3 covers the audio part and AVS Part 6 includes content creation. The AVS Part 2 encoding and decoding structures [8] are shown in Figure 6 (a) and (b).

System architecture

AVS Part 2 is hybrid coding based on spatial and temporal predictions, integer transform and entropy coding. The system architecture is illustrated in Figure 6 [8]

Intra Prediction

Spatial prediction as shown in Figure 7 is used in intra coding in AVS part 2 to exploit spatial correlations of picture. The intra prediction is based on 8x8 block. The intra prediction method is derived from the neighboring pixels in left and top blocks. There are five luminance intra
prediction modes, and four chrominance intra prediction modes. The reconstructed pixels of neighboring blocks before deblocking filter are used on reference pixels for the current block.

![Fig.7 Five different modes for 8 x8 block intra luminance prediction [8].](image)

**Deblocking filter**

The deblocking filter is used to reduce/eliminate the block artifacts and enhance both subjective and objective performance. AVS Part 2 deblocking filter first calculates the boundary strength (BS) of each block boundary, and then applies different filters for different BS.

**VLC**

AVS Part 2 utilizes an efficient context based 2D-VLC entropy coder for coding 8x8 block-size transform coefficients. 2D-VLC means that a pair Run-Level is regarded as one event and jointly coded.

**CONCLUSION**

This project aims at a thorough study, implementation and exhaustive comparison of video coding standards like H.264, Dirac pro and AVS part 2. Analysis will be carried out and different performance metrics like MSE, PSNR etc. will be evaluated for different high definition video sequences. Based on the values of these performance metrics, conclusions will be drawn as to which video coding standard is best suited for high definition video compression.

**References:**


[26] Dirac video codec – A programmer's guide: 
http://dirac.sourceforge.net/documentation/code/programmers_guide/toc.htm


