COMPARATIVE ANALYSIS OF DIRAC PRO-VC-2, H.264 AVC AND AVS CHINA-P7

A Thesis Submitted to the College of

Graduate Studies and Research

In Partial Fulfillment of the Requirements

For the Degree of Master of Science

In the Department of Electrical and Computer Engineering

University of Saskatchewan

Saskatoon, Saskatchewan

Canada

By

VISHESH KALRA

PERMISSION TO USE

ABSTRACT

Video codec compresses the input video source to reduce storage and transmission bandwidth requirements while maintaining the quality. It is an essential technology for applications, to name a few such as digital television, DVD-Video, mobile TV, videoconferencing and internet video streaming. There are different video codecs used in the industry today and understanding their operation to target certain video applications is the key to optimization. The latest advanced video codec standards have become of great importance in multimedia industries which provide cost-effective encoding and decoding of video and contribute for high compression and efficiency. Currently, H.264 AVC, AVS, and DIRAC are used in the industry to compress video. H.264 codec standard developed by the ITU-T Video Coding Experts Group (VCEG) together with the ISO/IEC Moving Picture Experts Group (MPEG). Audio-video coding standard (AVS) is a working group of audio and video coding standard in China. VC-2, also known as Dirac Pro developed by BBC, is a royalty free technology that anyone can use and has been standardized through the SMPTE as VC-2. H.264 AVC, Dirac Pro, Dirac and AVS-P2 are dedicated to High Definition Video, while AVS-P7 is to mobile video. Out of many standards, this work performs a comparative analysis for the H.264 AVC, DIRAC PRO/SMPTE-VC-2 and AVS-P7 standards in low bitrate region and high bitrate region. Bitrate control and constant QP are the methods which are employed for analysis. Evaluation parameters like Compression Ratio, PSNR and SSIM are used for quality comparison. Depending on target application and available bitrate, order of performance is mentioned to show the preferred codec.

ACKNOWLEDGEMENTS

I would like to express my gratitude to my supervisors, Dr. Anh v. Dinh and Dr. Khan A. Wahid, for introducing me to this research field and for their continuous support and guidance throughout the project. Their vast knowledge and expertise in this field added considerably to my graduate study experience.

I would also like to thank my thesis committee for their guidance and insightful comments. I would also like to acknowledge my financial support from NSERC. Finally, I would like to thank my family for their love, encouragement and support.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AVC	Advanced Video Coding
AVS	Audio Video Coding Standard
AVS-P2	Audio Video Coding Standard Part 2
AVS-P7	Audio Video Coding Standard Part 7
B	Bi-directional Predictive
BBC	British Broadcasting Corporation
BP	Baseline Profile
CABAC	Context Adaptive Binary Arithmetic Coding
CAVLC	Context Adaptive Variable Length Coding
CIF	Common Intermediate Format
CR	Compression Ratio
dB	Decibels
DCT	Discrete Cosine Transform
DVD	Digital Video Disc
EP	Extended Profile
HD	High Definition
HiP	High Profile
I	Intra
ICT	Integer Cosine Transform
IEC	International Electrotechnical Commission
ISO	International Standard Organization
ITU	International Telecommunication Union
MB	Macroblock
MC	Motion Compensation
ME	Motion Estimation
MV	Motion Vector
MP	Main Profile
MPEG	Moving Picture Effects Group
MSE	Mean Square Error

NAL	Network Abstraction Layer
P	Inter Predictive
PSNR	Peak Signal to Noise Ratio
PC	Personal Computer
SD	Standard Definition
SMPTE	Society of Motion Pictures and Television Engineers
SSIM	Structural Similarity Index
QP	Quantization Parameter
QF	Quality Factor
QCIF	Quarter Common Intermediate Format
VLC	Variable Length Coding

CHAPTER 1. INTRODUCTION TO VIDEO COMPRESSION

In this chapter, Video compression is defined. The need of video codec for efficient compression is highlighted. Motivation and objective are discussed in detail and parameters on basis of which three video codec standards analyzed are mentioned. This Chapter also outlines the thesis.

1.1 Video Compression

Broadcast television and home entertainment have been revolutionized by the advent of digital TV and DVD-video. These applications and many more are made possible by the standardization of video compression technology. Video compression is an essential technology for applications such as digital television, DVD-Video, mobile TV, videoconferencing and internet video streaming. Figure 1.1 below is an example of home media ecosystem in which video is transmitted among various devices. For fast transmission and quality reservation, video needs to be compressed. To start with, the size of a recorded video is large. A one second video recorded from a digital camcorder can have a size of more than 1 Mb (Megabit). Because it takes up so much space, video should be compressed before it is transferred. There are two kinds of compressions: lossy and lossless. Lossy compression means that the compressed file has less data than the original file. In some cases this translates to lower quality files, because information has been “lost”, hence the name. Lossless compression is exactly what it sounds like, compression where none of the information is lost [1].

Figure 1.1 An example of a home media ecosystem [1]

First we define codec which is a short version for encoder-decoder, the software that takes a raw data file and turns it into a compressed file. Because compressed files only contain some of the data found in the original file, the codec is the necessary "translator" that decides what data makes it in to the compressed version and what data gets discarded [1]. The latest advanced video codec standards have become of great importance in many industries which specialize in multimedia technologies. Standards provide cost-effective encoding and decoding of video and contribute for high compression and efficiency. Currently, the H.264 AVC, AVS and DIRAC are used in the industry to compress video. H.264 AVC 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) [2]. Audio-video coding standard (AVS) is a working group of audio and video coding standard in China [3]. VC-2, also known as Dirac Pro developed by BBC, is a royalty free technology that anyone can use and has been standardized through the SMPTE as VC-2 [4].

1.2 Motivation and Objective

Out of many previous works published, there are three prominent ones. The first one deals with the comparison between DIRAC, H.264 AVC and AVS-P2 [5] concluding that H.264 performs better than AVS and Dirac in general. The second one on the other hand has been related to the comparison between AVS-P2 and H.264 AVC [6], which shows that AVS has a good tradeoff between performance complexity for specific applications. The third one compares between DIRAC and H.264 AVC [7], which concludes that Dirac nearly equals H.264 in performance robust codec.

Dirac Pro/VC-2 is a comparatively new standard. Therefore, this work performs a comparative analysis for the H.264 AVC, DIRAC PRO/VC-2 and AVS-P7 standards which have not been previously published. The performance analysis used in this work is done by varying the Bitrate and QP (Quantization Parameter) for different video sequences including QCIF (Quarter-Common Intermediate Format), CIF (Common Intermediate Format), SD (Standard Definition) and HD (High Definition). Different profiles of the three codecs are also discussed and compared. The results are based on the Compression Ratio (CR), the peak signal to noise ratio (PSNR), and the Structural Similarity Index (SSIM). The SSIM is the main parameter for measuring the similarity between the original and the reconstructed sequences. Section 1.2.1 to Section 1.2.3 defines these parameters in detail.

1.2.1 Compression Ratio (CR)

The compression ratio depends on a variety of factors, such as video compressor, video quality, and video format. This ratio is used to quantify the reduction in representation size of the data produced by a compression algorithm. The compression ratio is defined as the size of the uncompressed video compared to that of the compressed video in the case of lossy video codecs. For example, if an uncompressed video is of size 15 Mb and a compressed of size 5 Mb, 3:1 is the compression ratio, in short 3.

1.2.2 Peak signal-to-noise ratio (PSNR)

Peak signal-to-noise ratio, often abbreviated as PSNR, is the ratio between the maximum possible power of a signal and the power of the corrupting noise that affects the quality of its representation. The PSNR is most commonly used as a measure of quality of the reconstruction of lossy compression codecs. When comparing codecs, PSNR is used as an approximation to human perception of reconstruction quality, therefore in some cases, one reconstruction may appear to be closer to the original than another, even though it has a lower PSNR (a higher PSNR would normally indicate that the reconstruction is of higher quality). PSNR and Mean Square Error (MSE) formulas are as follows [8]:

$$PSNR = 20 * \log10 (255 / \mathrm{sqrt}(MSE))$$

$$MSE = \frac{1}{MN} \sum_{y=1}^{M} \sum_{x=1}^{N} \left[I(x,y) - I'(x,y) \right]^2$$

I (x,y) is the original image and I' (x,y) is the reconstructed version, where as M, N are the dimensions of an image.

1.2.3 Structural Similarity Index (SSIM)

The structural similarity (SSIM) index is a method for measuring the similarity between two images. The SSIM index is a full reference metric, in other words, the measuring of image quality based on an initial uncompressed or distortion-free image as reference. SSIM close to 1 means reconstruction is almost identical to original video. It is designed to improve the traditional methods such as peak signal-to-noise ratio (PSNR) and mean squared error (MSE), which have proved to be inconsistent with human eye perception [9].

1.3 Thesis Outline

The first chapter is in brief an introduction to video compression and parameters used for analyzing the performance of a video codec. The second chapter describes different video codecs, i.e., H.264 AVC, Dirac Pro/SMPTE VC-2 and AVS China-P7. Features and architecture of codecs are also discussed. The third chapter relates to the setup associated with encoding video sequences, i.e., QCIF, CIF, SD, and HD using all the three video codecs mentioned above. The results, how the codecs fair against each other for the mentioned sequences are shown and analyzed in the fourth chapter. Finally the last chapter concludes the work and mentions for future exploration.

CHAPTER 2. VIDEO CODING STANDARDS

In this chapter, different coding standards, i.e., H.264 AVC, Dirac Pro/SMPTE-VC-2, and AVS-P7 are discussed. The encoder and decoder architectures are also briefly mentioned.

2.1 H.264 AVC

H.264 AVC is a codec standard capable of providing good video quality at substantially lower bitrates, promises better rate-distortion performance and compression efficiency than previous standards such as MPEG-2, H.263, or MPEG-4 Part 2 Visual without increasing complexity of the design [2]. Syntax specifications are simple and have the flexibility to be applied to a wide variety of applications such as video broadcasting, streaming and conferencing. Being a network friendly, H.264 AVC demonstrates a balance between coding efficiency and implementation complexity [10].

2.1.1 H.264 Codec

Simplified block diagrams of the H.264/AVC encoder and decoder are presented in Figure 2.1 and Figure 2.2 respectively. The encoder consists of three main functional units: prediction, core coding and entropy coding unit. The blocks mentioned in the Figure 2.1 are explained in brief from Section 2.1.1.1 to Section 2.1.1.5. In the H.264/AVC standard, each video frame is segmented into small blocks of pixel, called macroblocks (MB), and all the processing are performed on these macroblocks. Each macroblock consists of three components, Y, U, and V. Y is the luminance component, which represents the brightness information of the image. U and V are the chrominance components, and represent the colour information of the image. Figure 2.3 shows three components of an image with 4:2:0 subsampling which is known for encoding images by implementing less resolution for chroma information than for luma information, taking advantage of the human visual system's lower acuity for color differences than for luminance. The image is in YUV format, where Y represents the luminance component, and U, V represents the chrominance components. Y has the sample rate but U and V are each subsampled at a factor of 2 in both horizontally and vertically in the 4:2:0 subsampling.

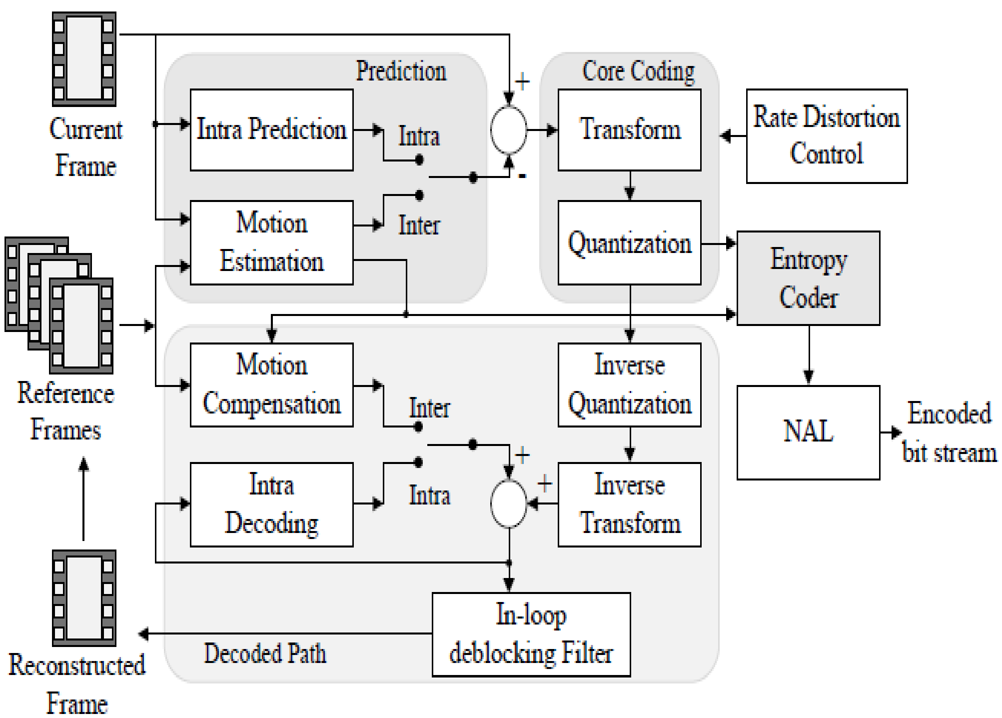


Figure 2.1 H.264 AVC Encoder [11]

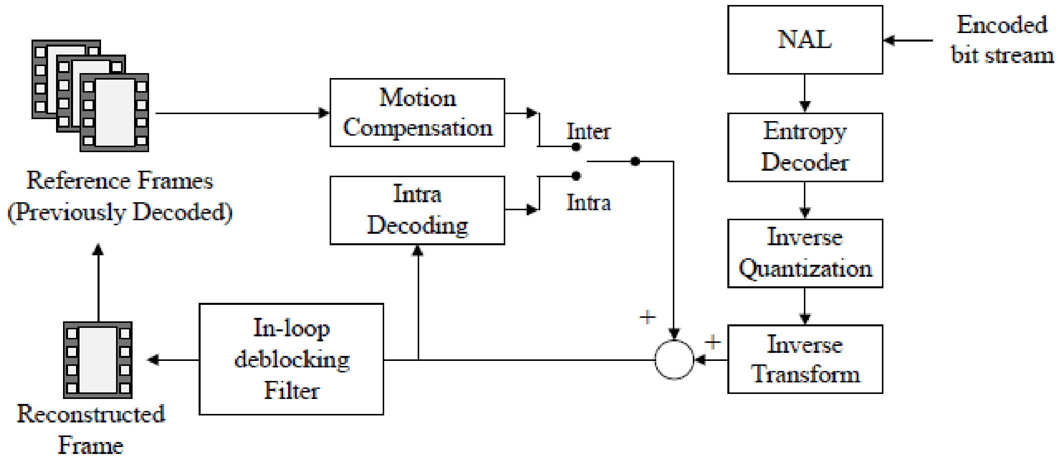


Figure 2.2 H.264 AVC Decoder [11]

2.1.1.1 Prediction Scheme

There are two important frames in Video Coding: I (intra) frames and P (predicted) frames. P frames contributes significantly towards the high compression ratios. One more frame predication method is also widely used, named B (bidirectional) frames. Intra frames are predicted using intra prediction which exploits the spatial correlation in each frame to reduce the amount of transmitted data also known as key frames or self-contained compressed images. Inter prediction predicts P and B frames. P frames are predicted using the previous P or I frame. Prediction is done by extracting the motion information from the frames and is normally encoded from the second frame onwards from the incoming frames. On the other hand, B-frames are bidirectional predicted frames. As the name suggests, B-frames rely on the frames preceding and following them. They can be encoded with lower quality without degrading the whole sequence. Since B-frames depend on both past and future frames, the decoder has to be fed with future I and P frames before decoder to process them [11].

(a) Original Video Frame

(b) Y Component

(c) U Component

(d) V Component

Figure 2.3 YUV Components of a Video Frame

2.1.1.2 Motion Estimation and Motion Compensation

The motion estimation algorithms are used in the encoding of Inter frames, i.e., P and B frames. H.264 encoding calculates MV (motion vectors), which means that the reference block is calculated by interpolation inside a block of real pixels. The motion vectors are different for luma and chroma blocks i.e., at quarter-pixel resolution and eighth-pixel accuracy respectively. Motion Compensation will decode the image that is encoded by Motion Estimation. This reconstruction of image is done from received motion vectors and the reference frame [11].

2.1.1.3 Transform and Quantization

Transform unit is used to compress the encoded Inter-frames or Intra-frames. The mostly used transform in H.264 is the Discrete Cosine Transform (DCT). Quantization block reduces the amount of information by dividing each coefficient by a particular index, i.e., Quantization Parameter (QP) to have further compression [11]. Smaller the QP, almost all the detail will be retained whereas at larger QP, quality distorts. QP of I, P, B frames is mentioned in Chapter three.

2.1.1.4 De-blocking Filter

It filters the reference frames from the frame buffer prior to use them in prediction which can significantly improve the objective and perceptual quality.

2.1.1.5 Entropy Coding and Network Abstraction Layer (NAL)

A lossless encoding block uses two types of coding such as context-adaptive variable-length coding (CAVLC) and Context-adaptive binary-arithmetic coding (CABAC) to convert non binary data to binary data by compression them further. With the knowledge of the probabilities of syntax elements in a given context, syntax elements in the video stream can be losslessly compressed by methods such as ex-golomb and arithmetic coding. All the compressed data is packetized in Network friendly format by NAL unit.

2.1.2 H.264 Profiles

The H.264 standard is a "family of standards", that include the following sets of capabilities, referred to as "profiles", targeting specific classes of applications [2] [10]. The various profiles are explained briefly below:

1. Baseline Profile (BP): It includes I and P slice coding only having enhanced error resilience and Context-adaptive variable-length coding (CAVLC). Applications include videoconferencing and mobile services.

2. Main Profile (MP): It is intended as the mainstream consumer profile for broadcast and storage applications. The Main profile includes all three I, P and B slices, interlaced coding and either CAVLC or Context-adaptive binary arithmetic coding (CABAC). Applications include standard-definition digital TV broadcasting.

3. Extended Profile (EP): The extended profile on the other hand is a superset of the baseline profile. It includes B slice, SP (switched prediction) slice and SI (switched intra) slices, data partitioning, and interlaced coding. Applications include streaming video.

4. High Profile (HiP): It supports the 8-bit video with 4:2:0 sampling. It uses B slices and CABAC. Applications include storage (blue-ray disc) and high definition TV broadcast.

The common coding parts for the profiles are listed below [10]:

1. I slice (Intra-coded slice)
2. P slice (Inter-Predictive coded slice)
3. Context-based Adaptive Variable Length Coding (CAVLC)

A brief description about the above has been done in Section 2.1.1.1 and Section 2.1.1.5.

2.2 Dirac Pro/SMPTE-VC-2

Dirac Pro is a version of the Dirac family of video compression tools, optimized for professional production and archiving applications, especially where the emphasis is on quality and low latency. Typical production processes require lossless or virtually lossless compression with low latency. Dirac has been streamlined to meet these requirements where as Dirac Pro is designed for simplicity, efficiency and speed. Dirac Pro is intended for high quality applications with lower compression ratios. Dirac Pro is an open technology, which will works on all of the major operating systems, such as Windows, Macintosh or Linux. As it is an open system, it is easy to import into a wide range of hardware, from specialized signal processors to application-specific integrated circuits [4].

2.2.1 Features of Dirac Pro

Dirac was the first codec developed by BBC (British Broadcasting Corporation). The main difference between Dirac and Dirac Pro is in the treatment of the final process in compression - the arithmetic coding. In addition to processing, Arithmetic coding also introduces delay. The arithmetic coding produces most efficiency savings with highly compressed material. There is little benefit to be gained with the low compression used in the top-end production. Dirac Pro therefore omits the arithmetic coding. Dirac Pro supports the following technical features, required by professional end-users [4] [12]:

- Intra-frame Prediction only (forward and backward prediction modes are also available if required). It includes self contained compressed I frames.
- Low complexity for decoding

- Open Specification
- Multiple vendors
- Support for multiple HD image formats and frame rates [12] [13].

Both Dirac and Dirac Pro are open Technologies and the Dirac software source code is licensed under the Mozilla Public License Version 1.1. Dirac Pro is being standardized by the SMPTE as “VC-2” and the standardization is virtually complete [12].

2.2.2 Dirac Encoder

The encoder has the architecture shown in Figure 2.4, whilst the decoder performs the inverse operations.

There are four main elements or modules of the encoder:

- Transform and scaling involve taking frame data and applying a transform (in this case the wavelet transform) and scaling the coefficients to perform quantization.
- Entropy coding is applied to the quantized transform coefficients and to motion vector (MV) data and performs lossless compression on them.
- Motion estimation (ME) involves finding matches for frame data from previously coded frames, trading off accuracy with motion vector Bitrate.
- Motion compensation (MC) involves using the motion vectors to predict the current frame, in such a way as to minimize the cost of encoding the residual data.

There are different profiles for VC-2 as listed in Table 2-1 and explained below:

- Low Delay Profile - Codes slices, no arithmetic coding, low latency (<10ms), low compression.
- Simple Profile - Codes whole pictures, no arithmetic coding-Higher compression, latency approximately equal to1 picture.
- Main Profile - Simple profile + arithmetic coding-Higher compression, more complex, more latency.

Table 2-1 Various Profiles of DIRAC PRO/VC-2

Profiles	**Low Delay**	**Simple**	**Main**
Complexity	Low	Low	Medium
Coding Units	Slices	Pictures	Pictures
Latency	Very low	1-2 Pictures	2-3 Pictures
Arithmetic Coding	No	No	Yes

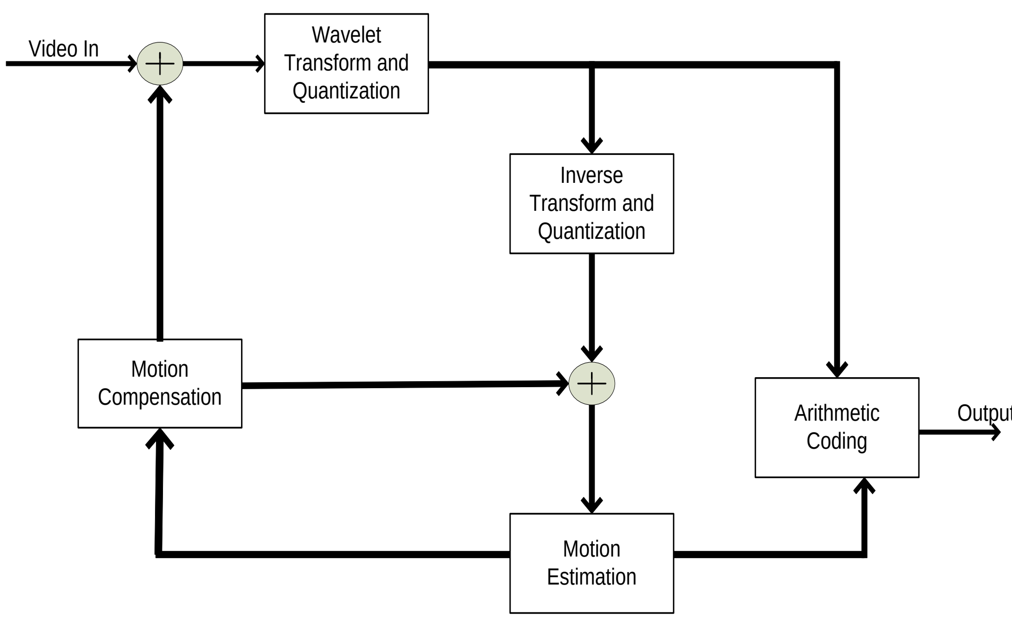


Figure 2.4 Dirac Encoder

2.3 AVS CHINA-P7

Audio-video coding standard (AVS) is a working group of audio and video coding standard in China, which was established in 2002 [14]. AVS-P7 standard is the seventh part of the standards developed by the Audio Video Coding Standard (AVS) Workgroup of China also known as Jiben Profile. There are 10 parts of the AVS standard, as shown in Table 2-2 [14].

Table 2-2 Different Parts of the AVS-China

Parts of AVS	Category
Part 1	System
Part 2	Video
Part 3	Audio
Part 4	Conformance Test
Part 5	Reference Software
Part 6	Digital Media Rights Management
Part 7	Mobile Video
Part 8	Transmit AVS via IP Network
Part 9	AVS File Format
Part 10	Mobile Speech and Audio Coding

2.3.1 AVS-P7 Codec

The encoder and decoder block diagrams of the AVS-P7 are shown in Figures 2.5 and 2.6, respectively. In terms of modules, there is a similarity between AVS-P7 and H.264. Each input macroblock needs to be predicted (intra predicted or inter predicted) in AVS-P7. In an AVS-P7 encoder, the S_0 (switch) is used to select the proper prediction method for the current macroblock. In an AVS-P7 decoder, the S_0 is controlled by the macroblock type of the current macroblock. The intra predictions are derived from the neighboring pixels in the left and top blocks with respect to original pixel. The unit size of intra prediction is 4×4 because of the 4×4 integer cosine transform (ICT) used by AVS-P7.

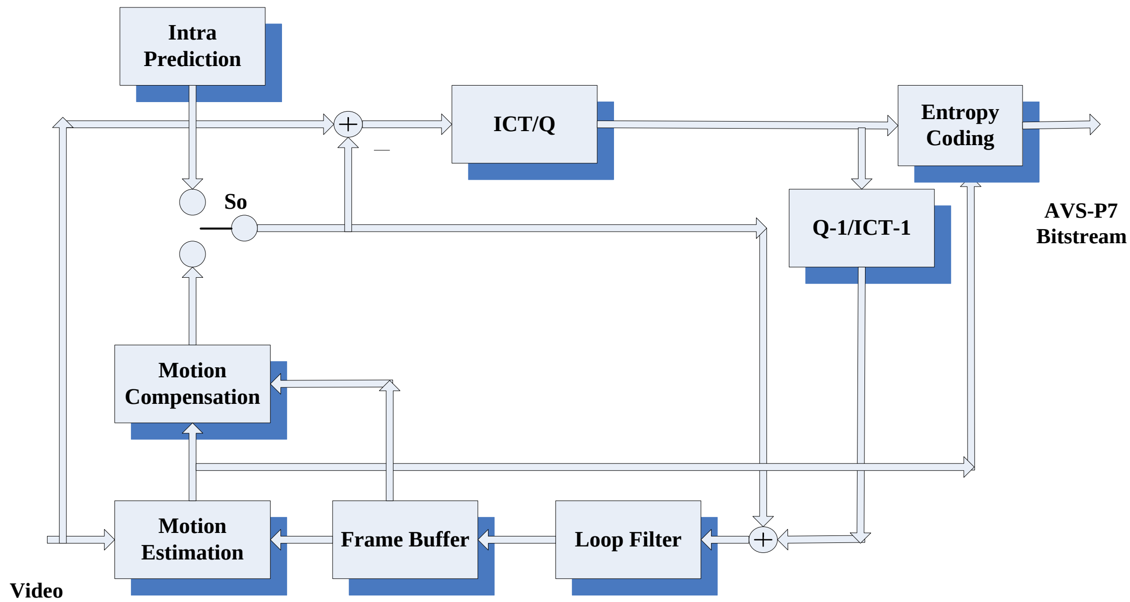


Figure 2.5 AVS-P7 Encoder

The inter predictions are derived from the decoded frames. Seven types of block sizes, i.e., 16×16, 16×8, 8×16, 8×8, 8×4, 4×8, and 4×4 are supported in AVS-P7. The precision of motion vector in inter prediction is up to 1/4 pixel [15]. The prediction residues are transformed with 4×4 ICT. The ICT coefficients are quantized using a scale quantizer. The scanning order of the quantized coefficients used in AVS-P7 is still zig-zag similar to that used in MPEG-1 and MPEG-2. AVS-M employs an adaptive VLC (Variable Length Coding) coding technique. The sum of prediction and current reconstructed error image form the reconstructed reference. AVS-P7 uses the deblocking filter in motion compensation loop. The deblocking process directly acts on the reconstructed reference first across the vertical edges and then across the horizontal edges. Obviously, different image regions and different Bitrates need different smoothes. Therefore, the deblocking filter is automatically adjusted in AVS-P7 depending on activities of the blocks and QP parameters. AVS-P7 only supports progressive video sequence. So in AVS-P7, one picture is one frame. AVS-P7 supports the 4:2:0 format. AVS-P7 specifies two types of pictures, which is I picture and P picture. In AVS-P7, P picture can have a maximum of two reference frames for forward prediction [15].

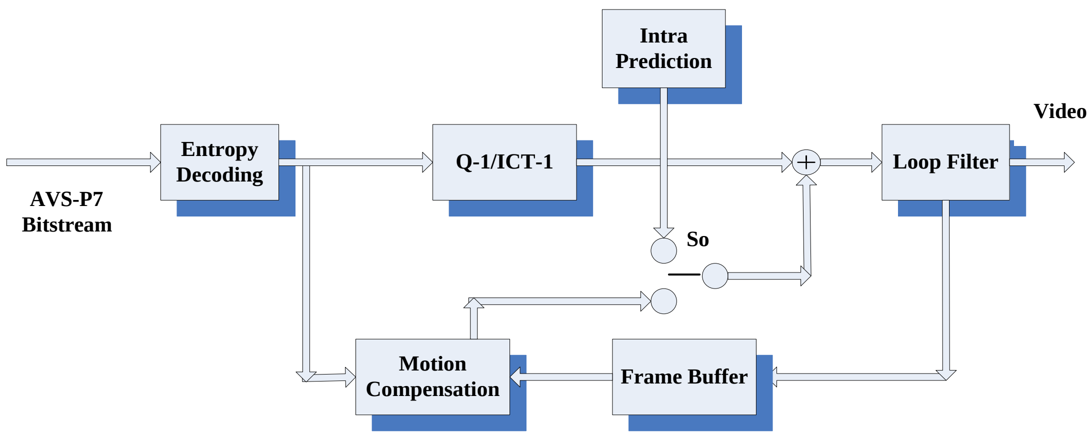


Figure 2.6 AVS-P7 Decoder

2.4 Parametric Comparison between Different Standards

Table 2-3 below highlights various features on which the three codecs differ. It can be said that Dirac Pro and AVS-P7 are less complex as both do not support B frames. Dirac Pro is the least complex out of the three because it does not support P frames either and also does not have a De-blocking filter. Dirac Pro is an intra frame codec which follows CABAC only. Other difference is, it employs wavelet transform as compared to DCT

Transform used in H.264 and AVS-P7. It becomes quite important to study how these codecs compare with each other.

Table 2-3 Architecture Comparison [5]

Algorithmic Element	**MPEG-4 AVC (H.264)**	**Dirac Pro**	**AVS China Part 7**
Intra Prediction	4x4 spatial 16x16 spatial I-PCM	4x4 Spatial (forward, backward)	Intra_4x4 (4x4 spatial). Direct Intra Prediction
Picture coding type	Frame Field Picture AFF MB AFF	Intra – Frame, Field (Interlace, Progressive)	Frame
Motion compensation block size	16×16, 16×8, 8×16, 8×8, 8×4, 4×8, 4×4	N/A	16×16, 16×8, 8×16, 8×8, 8×4, 4×8
Motion vector Precision	Full pel Half pel Quarter pel	N/A	1/4 pel
P frame type	Single reference Multiple reference	No P frames	Single and multiple reference (maximum of 2 reference frames)
B frame type	One reference each way, Multiple reference, Direct & spatial direct weighted prediction.	No B frames	No B frames.
In loop filters	De-blocking	None	De-blocking filter.
Entropy coding	CAVLC,CABAC	Context based adaptive binary arithmetic coding, Exp-Golomb coding.	Context based adaptive 2D variable length coding.
Transform	4×4 integer DCT 8×8 integer DCT	4×4 wavelet transform	4×4 DCT
Other	Quantization scaling matrices.	Quantization scaling matrices.	Quantization scaling matrices.

CHAPTER 3. COMPARISON METHODOLOGY

In order to encode a video, a proper procedure has to be followed and this chapter deals with the parameter settings of the reference software and basis on which analysis is done. Two regions are defined, i.e., low bitrate region (10 Kbps-8 Mbps) which highlights the encoding of QCIF, CIF and SD video to compare all the three standards H.264, AVS-P7 and Dirac Pro/VC-2 and high bitrate Region (10-20 Mbps) for encoding of High Definition Video to compare only H.264 and Dirac Pro/VC-2 since AVS-P7 is limited to SD Video. Low bitrate region and high bitrate region separates AVS-P7 with H.264 and Dirac Pro/VC-2. Section 3.1 explains the work flow.

3.1 Work Flow

Figure 3.1 shows the software work flow. Video sequences such as QCIF, CIF, SD and HD are used as shown in Figure 3.2. For example, H.264 JM reference software [16] encodes the original video into a compressed video format, i.e., xxx.264. Compressed video can be used to calculate compression ratio directly. Consequently, the xxx.264 is passed to the decoder to reconstruct the video sequence. Original video sequence and reconstructed video sequence help to generate PSNR and SSIM plots. The same procedure is followed for softwares of Dirac Pro/VC-2 (Shroedinger) [17] and AVS-P7 (Jiben) [18]. The comparison between H.264, Dirac Pro/VC-2 and AVS-P7 is divided into three different parts: encoding test sequences by varying bitrates for popular profiles, encoding test sequences by varying QP and impact of bitrate on all different profiles of codecs. For Section 3.1.1 and Section 3.1.2, H.264 High, VC-2 Main and AVS-P7 are chosen for analyses since these are the best profiles available for the respective codecs and for Section 3.1.3, all profiles have been considered.

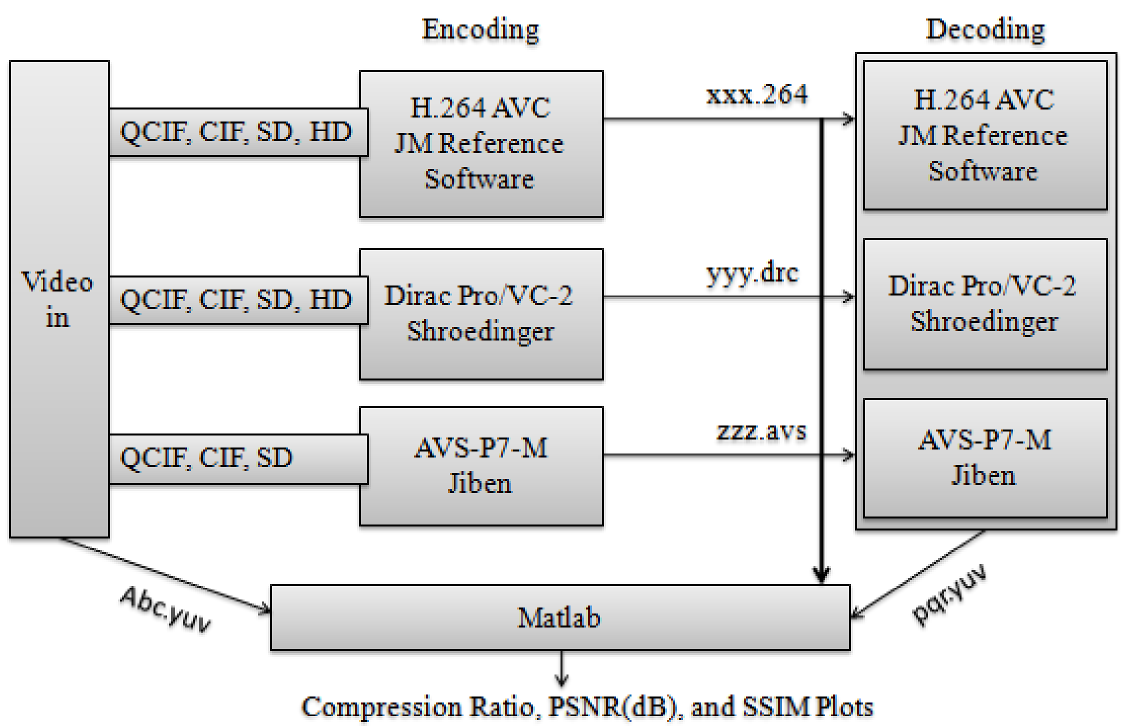


Figure 3.1 Software Flow for H.264, Dirac Pro/VC-2 and AVS-P7

3.1.1 Encoding test sequences by varying Bitrates

Constant bitrate is applied to all the three codecs. This is known as rate–distortion optimization which is a method of improving video quality during video compression. It is used by video encoders to decide what affect both file size and quality simultaneously. Performance was evaluated for three different parameters: Compression ratio, Peak-to-Peak Signal-to-Noise ratio (PSNR-Y), SSIM (Structural Similarity Index). The JM software for H.264, the Schroedinger for Dirac Pro/VC-2 and AVS-M (Jiben) for AVS-P7 were used to encode and decode the sequences. Encoder setup for QCIF and CIF

Sequence are shown in Table 3-1 and Table 3-4. Setup for SD is shown in Table 3-2. Table 3-3 and Table 3-5 highlight the setup for HD.

As far as JM Software is concerned, Intra Period is 0, which tells only the first frame is I frame (Original Frame). The same setting is kept for AVS-P7. Ideally, QP-I, B, P in H.264 JM are kept at 28, 30, 28. Since AVS-P7 follows prediction with respect to reference frames, QP-First Frame and QP-Remaining Frame make more sense. The bitrate is also kept the same for both to analyze how the codecs fair against each other. In Dirac Pro Shroedinger, Bitrate and QF (Quality Factor) are two main parameters on which analyses can be done. For Section 3.1.1 and Section 3.1.3, bitrate is kept the same and for Section 3.1.2, JM and AVS-P7 employs the same QP and in Dirac Pro, the QF is varied as discussed.

Table 3-1 Encoder Setup for H.264 AVC and AVS-P7 for QCIF and CIF Sequences

Simulation Setup	H.264 JM	AVS-P7
Video Sequence Used	Miss-America (QCIF), Stefan (CIF)	
Format	QCIF	CIF
Frame Size	176x144	352x288
Frame Rate	25	
Chroma Format	4:2:0	
Profile	High-Main-Extended-Baseline	Jiben
Intra Period	0	0
Number of B Frames	1	N.A.
QP-I	28	N.A.
QP-P	28	N.A.
QP-B	30	N.A.
QP-First Frame	N.A.	40
QP-Remaining Frame	N.A.	28
Rate Control	1	1
Bitrate (Kbps)	10-200 Kbps	10-200 Kbps

Table 3-2 Encoder Setup for H.264 AVC and AVS-P7 for QCIF and CIF Sequences

Simulation Setup	H.264 JM	AVS-P7
Video Sequence Used	Ice (SD)	
Format	SD	
Frame Size	704x576	
Frame Rate	30	
Chroma Format	4:2:0	
Profile	High-Main-Extended-Baseline	Jiben
Intra Period	0	0
Number of B Frames	1	N.A.
QP-I	28	N.A.
QP-P	28	N.A.
QP-B	30	N.A.
QP-First Frame	N.A.	40
QP-Remaining Frame	N.A.	28
Rate Control	1	1
Bitrate (Mbps)	1-8 Mbps	

Table 3-3 Encoder Setup for H.264 AVC and AVS-P7 for HD Sequence

Simulation Setup	H.264 JM
Video Sequence Used	Stockholm (HD)
Format	HD
Frame Size	1280x720
Frame Rate	30
Chroma Format	4:2:0
Profile	High-Main-Extended-Baseline
Intra Period	0
Number of B Frames	1
QP-I	28
QP-P	28
QP-B	30
QP-First Frame	N.A.
QP-Remaining Frame	N.A.
Rate Control	1
Bitrate (Mbps)	10-20 Mbps

Table 3-4 Encoder Setup for Dirac Pro for QCIF and CIF Sequences

Simulation Setup	Dirac Pro/VC-2 Schroedinger	
Video Sequence Used	Miss-America/Akiyo	Stefan/Bus
Format	QCIF	CIF
Frame Size	176x144	352x288
Frame Rate	25	
Chroma Format	4:2:0	
Profile	VC-2 Simple, VC-2 Main	
QF (Quality Factor)	0-10	
Bitrate (Kbps)	10-200 Kbps	

Table 3-5 Encoder Setup for Dirac Pro for SD and HD Sequences

Simulation Setup	Dirac Pro/VC-2 Schroedinger	
Video Sequence Used	Ice	Stockholm
Format	SD	HD
Frame Size	704X576	1280X720
Frame Rate	30	
Chroma Format	4:2:0	
Profile	VC-2 Simple, VC-2 Main	
QF (Quality Factor)	0-10	
Bitrate (Mbps)	1-8 Mbps	10-20 Mbps

3.1.2 Encoding test sequences by varying QP

Next, the comparison was performed by varying the quantization parameter, QP. The parameter 'QP' is present in the reference software of H.264 AVC, but does not exist in Dirac Pro/VC-2. Hence Quality Factor (QF), of the encoded video streams which is the amount of bits for each pixel in an encoded video, being present in Dirac Pro/VC-2 was found to be inversely proportional to the QP. Low QF means higher value of QP and vice-versa [19]. The range of QP in H.264 AVC is 0-51 [20] and the range of QF in VC-2 is 0-10 [21]. To have an analysis based upon the QP for both codecs, it was assumed that the maximum value of QF in VC-2 (i.e., 10) being the minimum QP (i.e., 0) of the H.264

AVC. Table 3-6 to Table 3-8 show the setup for QCIF – CIF, SD and HD sequences respectively.

Table 3-6 Encoder Setup for H.264 AVC and AVS-P7 for QCIF and CIF Sequences

Simulation Setup	H.264 JM	AVS-P7
Video Sequence Used	Akiyo (QCIF), Bus (CIF)	
Format	QCIF	CIF
Frame Size	176x144	352x288
Frame Rate	25	
Chroma Format	4:2:0	
Profile	High-Main-Extended-Baseline	Jiben
Intra Period	0	0
Number of B Frames	1	N.A.
QP-I	0-50	N.A.
QP-P	0-50	N.A.
QP-B	0-50	N.A.
QP-First Frame	N.A.	0-50
QP-Remaining Frame	N.A.	0-50
Rate Control	0	0

Table 3-7 Encoder Setup for H.264 AVC and AVS-P7 for SD Sequence

Simulation Setup	H.264 JM	AVS-P7
Video Sequence Used	Ice (SD)	
Format	SD	
Frame Size	704x576	
Frame Rate	30	
Chroma Format	4:2:0	
Profile	High-Main-Extended-Baseline	Jiben
Intra Period	0	0
Number of B Frames	1	N.A.
QP-I	0-50	N.A.
QP-P	0-50	N.A.
QP-B	0-50	N.A.
QP-First Frame	N.A.	0-50
QP-Remaining Frame	N.A.	0-50
Rate Control	0	0

Table 3-8 Encoder Setup for H.264 AVC for HD Sequence

Simulation Setup	H.264 JM
Video Sequence Used	Stockholm (HD)
Format	HD
Frame Size	1280x720
Frame Rate	30
Chroma Format	4:2:0
Profile	High-Main-Extended-Baseline
Intra Period	0
Number of B Frames	1
QP-I	0-50
QP-P	0-50
QP-B	0-50
QP-First Frame	N.A.
QP-Remaining Frame	N.A.
Rate Control	0

3.1.3 Impact of Bitrate on different Profiles of Codecs

In the third step, comparison between different profiles was performed with respect to the above mentioned parameters for QCIF sequence and HD sequence. The CIF and SD sequences were also included in the performance tests. The five original test sequences used for evaluation shown in Figure 3.2 are: “Miss-America” QCIF (176×144) [22], “Akiyo” QCIF (176×144) [22], “Stefan” CIF (352×288) [22], “Bus” CIF (352×288) [22], “Ice" standard-definition (SD) (704×576) [23] and “Stockholm" High-definition (HD) (1280×720) [23]. These sequences are chosen because they have been used in previous benchmark studies as mentioned in Section 1.2.

Profiles for which the video is encoded are mentioned in Table 3-1 to Table 3-8. Setup for the encoder is followed the same way as in Section 3.1.1.

(a) Miss-America QCIF (176x144)

(b) Akiyo QCIF (176x144)

(c) Stefan CIF (352x288)

(d) Bus CIF (352x288)

(e) Ice SD (704x576)

(f) Stockholm HD (1280x720)

Figure 3.2 Test sequences used in the comparison

CHAPTER 4. RESULTS

The results are classified as in low bitrate region and high bitrate region. Low bitrate region deals with QCIF, CIF and SD sequences and compares H.264, AVS-P7 and VC-2. High bitrate region deals with HD sequence and compares H.264 and VC-2 only as AVS-P7 is limited up to SD. In the last part, i.e., summary, all the results are tabulated for all the sequences to highlight the codecs with respect to their order of performance.

4.1 Introduction

This section discusses the results in form of various plots for various sequences like QCIF, CIF, SD and HD Video, classified as in Low bitrate Region (1-8 Mbps) and high bitrate region (10-20 Mbps).

4.2 Low Bitrate Region

This section deals with analyzing all the three codecs for QCIF and CIF sequences with respect to bitrate, because all the three codecs support these sequences in the 10 Kbps-8 Mbps region. Sequences are Miss America for QCIF and Stefan for CIF and encoding them to find out how a fairly good visual quality can be obtained around the suggested bitrate.

For QCIF and CIF, bitrate is varied from 10-200 Kbps as mentioned in Section 3.1.1. Results of SD video (low bitrate region) are summarized along with other results in Section 4.7.

4.2.1 Miss-America QCIF Sequence

This section deals with encoding of the Miss America QCIF sequence for all the three codecs with respect to bitrate from 10-100 Kbps and analyzing them in terms of Compression Ratio, PSNR and SSIM.

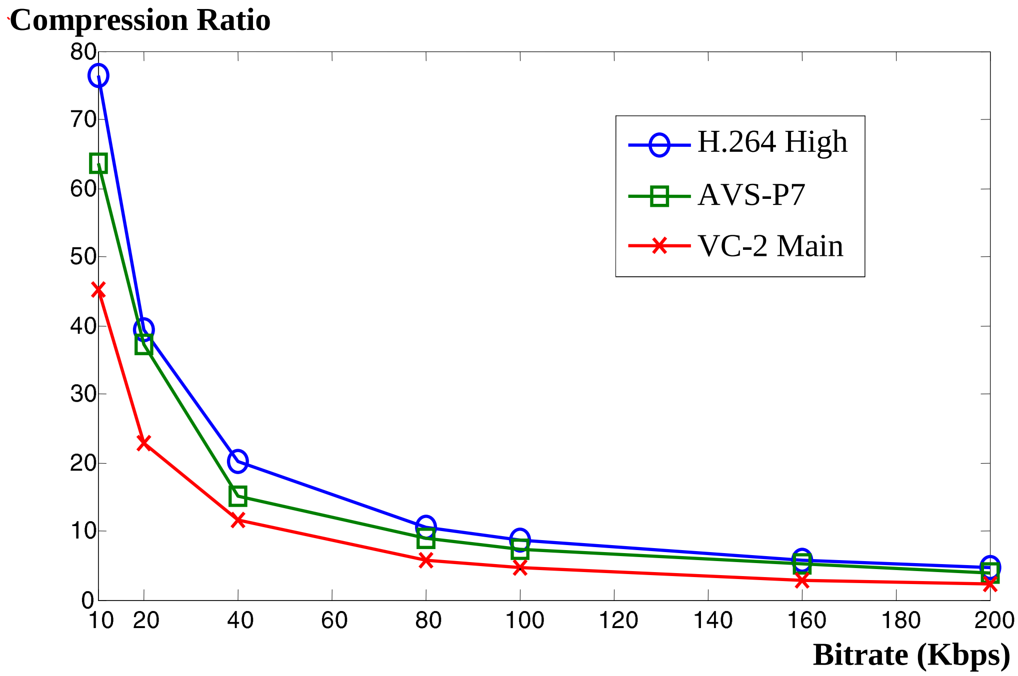


Figure 4.1 Compression Ratio of H.264 High, AVS-P7 and Dirac Pro/VC-2 Main for "Miss-America" sequence

Encoding for the QCIF in JM and AVS-P7 is done according to Table 3-1 from Chapter 3, where as for VC-2, Table 3-4 is followed. H.264 High, VC-2 Main and AVS-P7 are used as profiles for this section. Figure 4.1 shows the Compression Ratio vs Bitrate (Kbps) curves for QCIF (Miss-America). The curve is non-linear. The relationship is exponential, the compression ratio decreases as the bitrate increases. H.264 High performs better than AVS-P7 and VC-2 Main. But at bitrates close to 200 Kbps, the compression converges to a similar value. Normally, video is encoded at a rate which will have a suitable compression instead of having a negligible compression. The bitrate from 60-100 Kbps has a good compression ratio for all the three codecs, highest being of H.264.

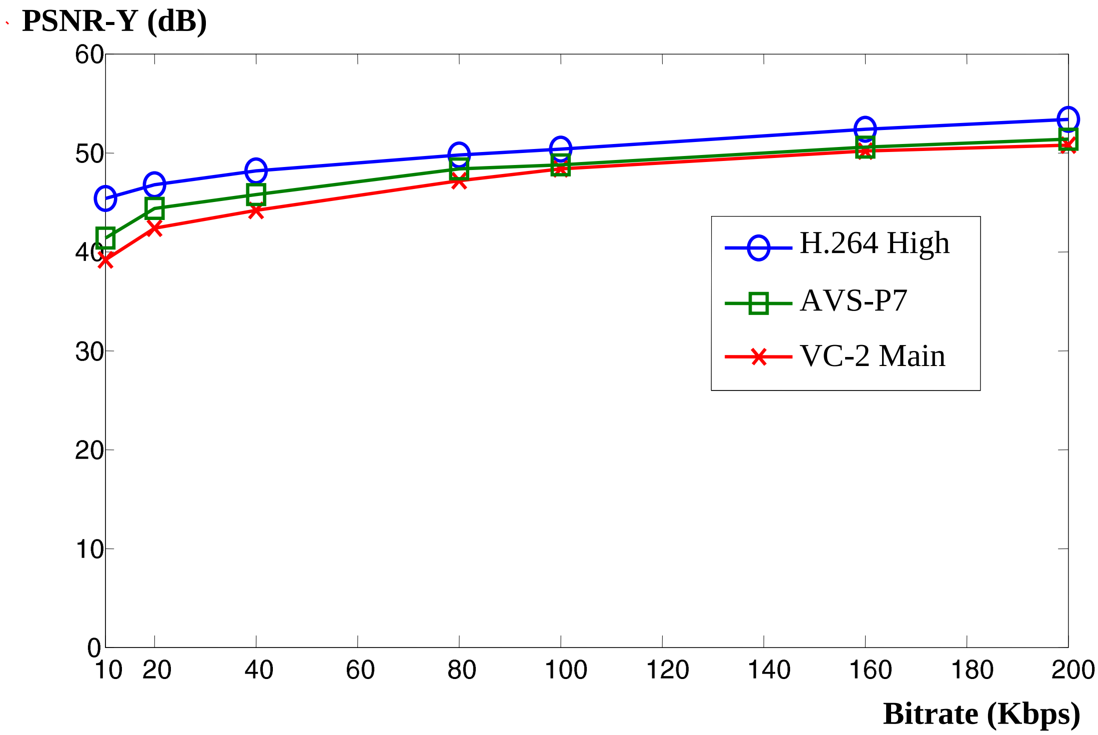


Figure 4.2 PSNR-Y of H.264 High, AVS-P7 and Dirac Pro/VC-2 Main for "Miss-America" sequence

Figure 4.2 shows the PSNR-Y (dB) vs Bitrate (Kbps) curves for QCIF (Miss-America). H.264 outperforms AVS-P7 and VC-2. While VC-2 and AVS-P7 converge to a same value around 100 Kbps, PSNR-Y (dB) becomes comparable. So the clear winner in terms of PSNR-Y is H.264 for QCIF sequence. Assumption can be made for bitrates greater than 200 Kbps that codecs tend to have approximately same PSNR. Emphasis is paid on high compression and high PSNR-Y which is noticeable in the range of 60-100 Kbps.

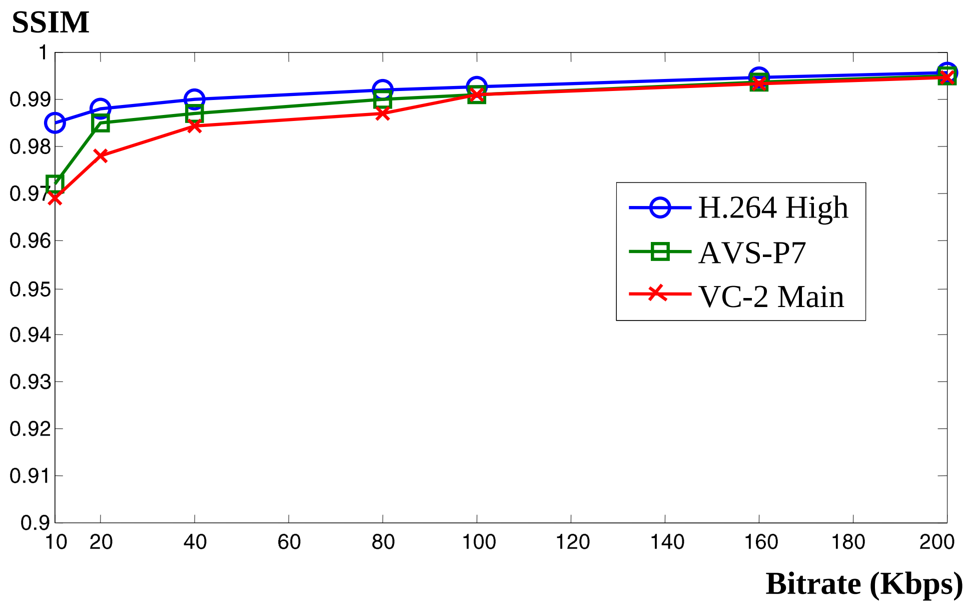


Figure 4.3 SSIM of H.264 High, AVS-P7 and Dirac Pro/VC-2 Main for "Miss-America" sequence

Figures 4.3 shows the SSIM vs Bitrate (Kbps) curves for QCIF (Miss-America). There is an increasing trend for AVS-P7 and VC-2 until 20 Kbps but saturates after that. The H.264 codec has a better SSIM until 100 Kbps but after that all the three standards saturate to a value close to 1. Figure 4.4 shows the reconstructed video sequences of QCIF at 100 Kbps. At 100 Kbps, since all codecs have suitable SSIM (close to 1), good reconstruction is possible in this case.

(a) H.264 (100 Kbps) (b) VC-2 (100 Kbps) (c) AVS-P7 (100 Kbps)

Figure 4.4 Reconstructed video sequences at 100 Kbps

4.2.2 Stefan CIF Sequence

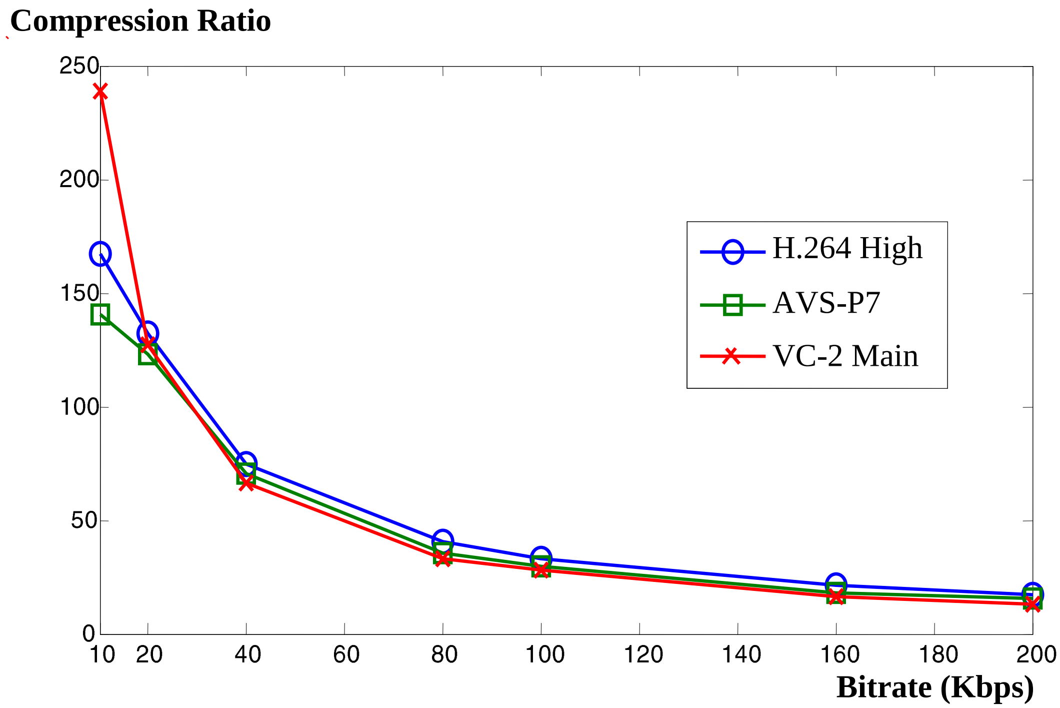


Figure 4.5 Compression Ratio of H.264 High, AVS-P7 and Dirac Pro/VC-2 Main for "Stefan" sequence

Figure 4.5 shows the Compression Ratio vs Bitrate (Kbps) curves for CIF (Stefan). The curves in this case follow the same trend as QCIF sequence, i.e., H.264 High > AVS-P7 > VC-2 Main. VC-2 Main is trailing behind, but around 200 Kbps, it is so similar to other two codecs. Better compression is possible in range of 80-100 Kbps, H.264 performing better but around that range we get a good PSNR-Y too as shown in the Figure 4.6. So for target applications, we get a glimpse of a range where one can get efficient compression as well as a good quality reconstruction for different codecs. Suitable applications for this range are video streaming over cellphones or multimedia messaging service, i.e., receiving video messages.

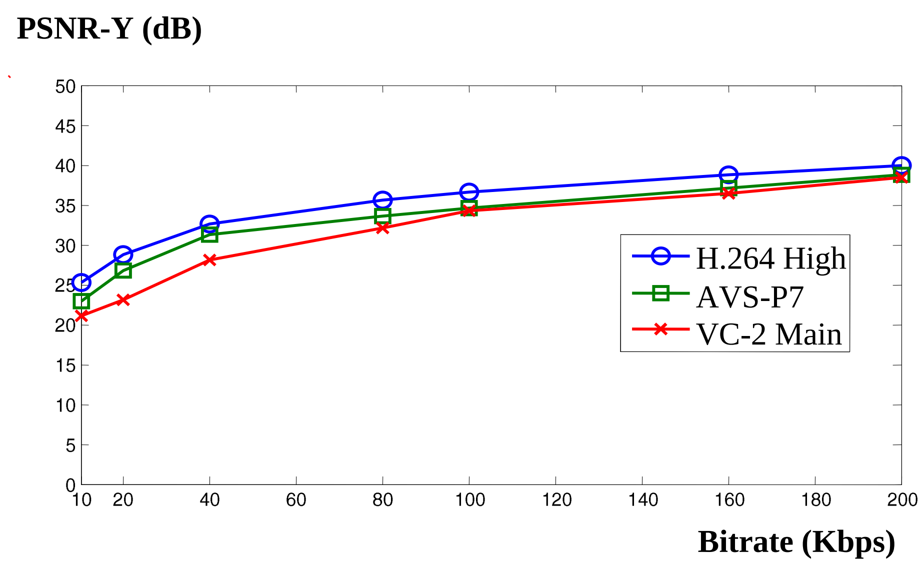


Figure 4.6 PSNR-Y of H.264 High, AVS-P7 and Dirac Pro/VC-2 Main for "Stefan" sequence

Figure 4.6 shows the PSNR-Y (dB) vs Bitrate (Kbps) curves for CIF (Stefan). The trend in both cases (QCIF and CIF-PSNR-Y) is found similar with H.264 > AVS-P7 > VC-2. Around 100 Kbps, it is found that AVS-P7 is comparable to VC-2.

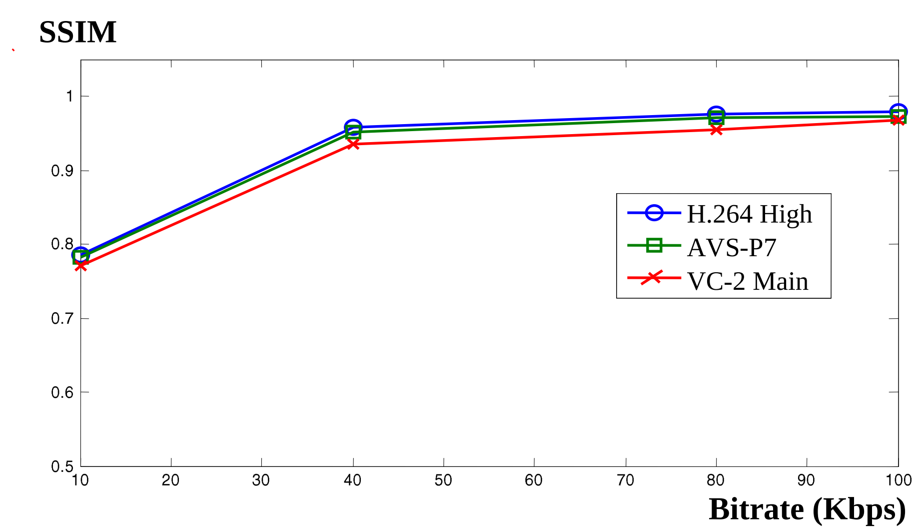


Figure 4.7 SSIM of H.264 High, AVS-P7 and Dirac Pro/VC-2 Main for "Stefan" sequence

Figure 4.7 shows the SSIM vs Bitrate (Kbps) curves for QCIF (Miss-America) and CIF (Stefan). As far as CIF Sequence is concerned, the trend is increasing for all the three standards until 40 Kbps and then it stops incrementing as before and the SSIM value comes close to 1 for all the three around 100 Kbps. Figure 4.8 shows the reconstructed video sequences of CIF at 100 Kbps for different codecs. It can be observed that H.264 performs much better.

(a) H.264 (100 Kbps) (b) VC-2 (100 Kbps) (c) AVS-P7 (100 Kbps)

Figure 4.8 Reconstructed video sequences at 100 Kbps.

4.3 Impact of QP-Low Bitrate Region

Low bitrate region corresponds to a region where we compare all the three codecs, H.264, AVS-P7 and VC-2 which encodes QCIF, CIF and SD sequence.

The sub-sections deals with encoding of the Akiyo QCIF and Bus CIF for all the three codecs with respect to QP (0-50) and analyzing them in terms of Compression Ratio, PSNR and SSIM around the region QP (20-30), where fairly good quality can be achieved because of the fact that at much higher QP, more spatial detail is removed and quality distorts.

4.3.1 Akiyo QCIF Sequence

Compression Ratio

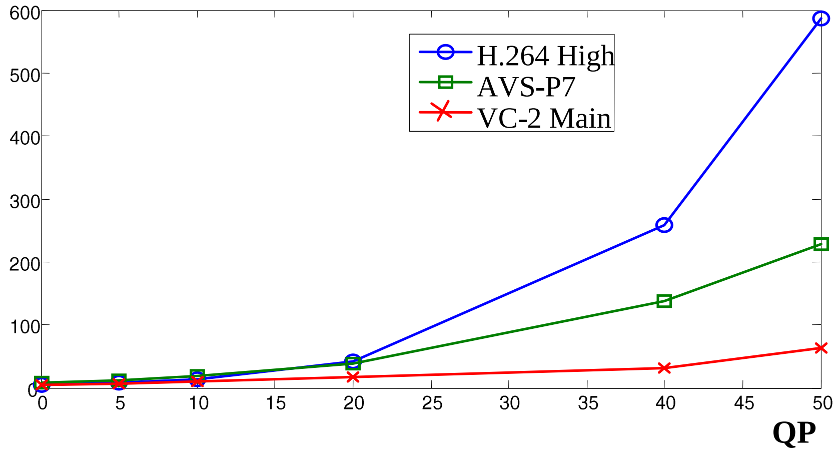


Figure 4.9 Compression Ratio of H.264 High, AVS-P7 and Dirac Pro/VC-2 Main for "Akiyo" sequence

Figure 4.9 shows the Compression Ratio vs QP (Quantization Parameter) curves for QCIF (Akiyo). Till QP=20, the compression ratio is quite low and almost similar for all of the three codecs, but from 20 to 50, with H.264 performing better than the other two.

PSNR-Y (dB)

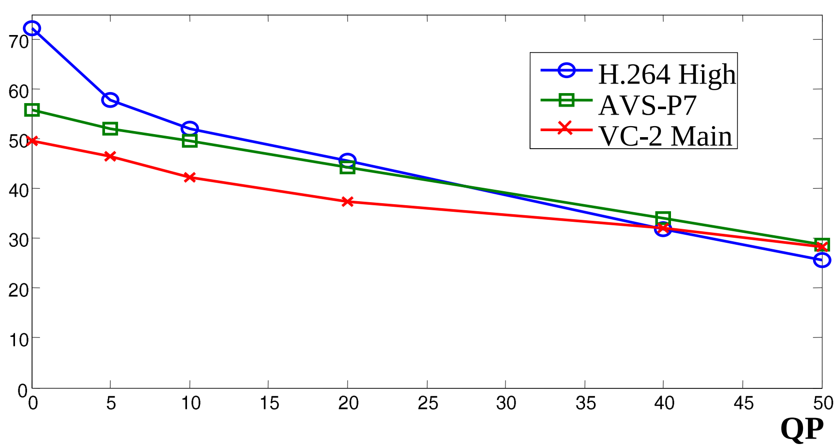


Figure 4.10 PSNR-Y of H.264 High, AVS-P7 and Dirac Pro/VC-2 Main for "Akiyo" sequence

Figure 4.10 shows PSNR-Y (dB) vs QP (Quantization Parameter) curves for QCIF sequence. From QP=0 to QP=20, the order of performance is H.264>AVS-P7>VC-2. At QP=25, AVS-P7 overtakes H.264 and VC-2 overtakes H.264 at QP=45 and almost equal to AVS-P7 at QP=50. It is evident that at higher QP, H.264 performance decreases as compared to other codecs.

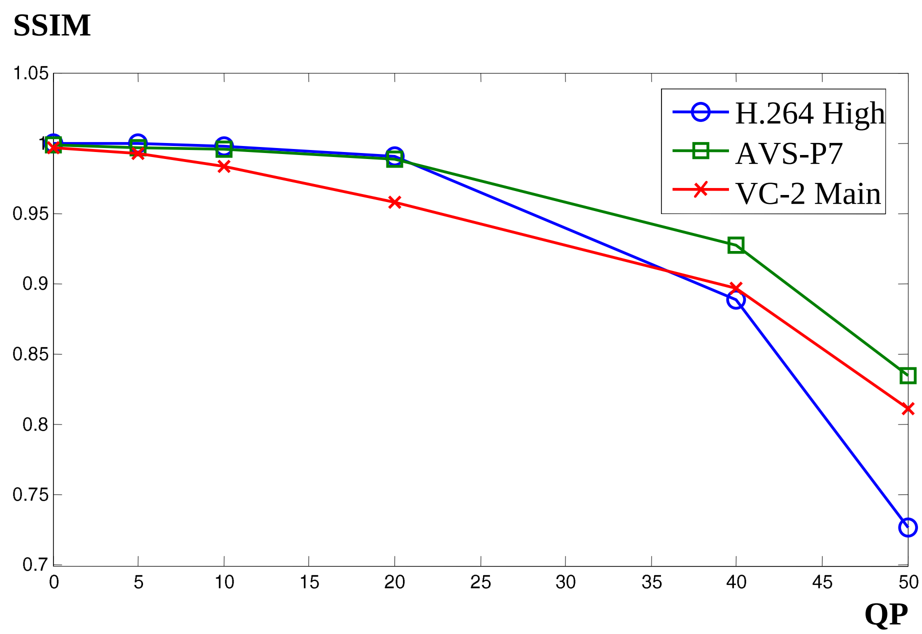


Figure 4.11 SSIM of H.264 High, AVS-P7 and Dirac Pro/VC-2 Main for "Akiyo" sequence

Figure 4.11 shows the SSIM vs QP (Quantization Parameter) curves for QCIF (Akiyo) sequence. Since QP in H.264 is inversely proportional to the Quality in Dirac Pro, decrease in SSIM as QP increases is no exception. For QCIF SSIM is approximately equal for H.264 and AVS-P7 uptill QP=20 but better than VC-2. At QP=40, the trend reverses and the order goes as follows, AVS-P7 > VC-2 > H.264. Figure 4.12 shows the reconstructed video sequences of QCIF at QP=0 and QP=50. QP=0 has no artifacts for all the three codecs but at QP=50, AVS-P7 > VC-2 > H.264, in terms of quality.

(a) H.264 (QP=0)　(b) AVS-P7 (QP=0)　(c) VC-2 (QP=0)

(d) H.264 (QP=50)　(e) AVS-P7 (QP=50)　(f) VC-2 (QP=50)

Figure 4.12 Reconstructed sequences for varying QP

4.3.2 Bus CIF Sequence

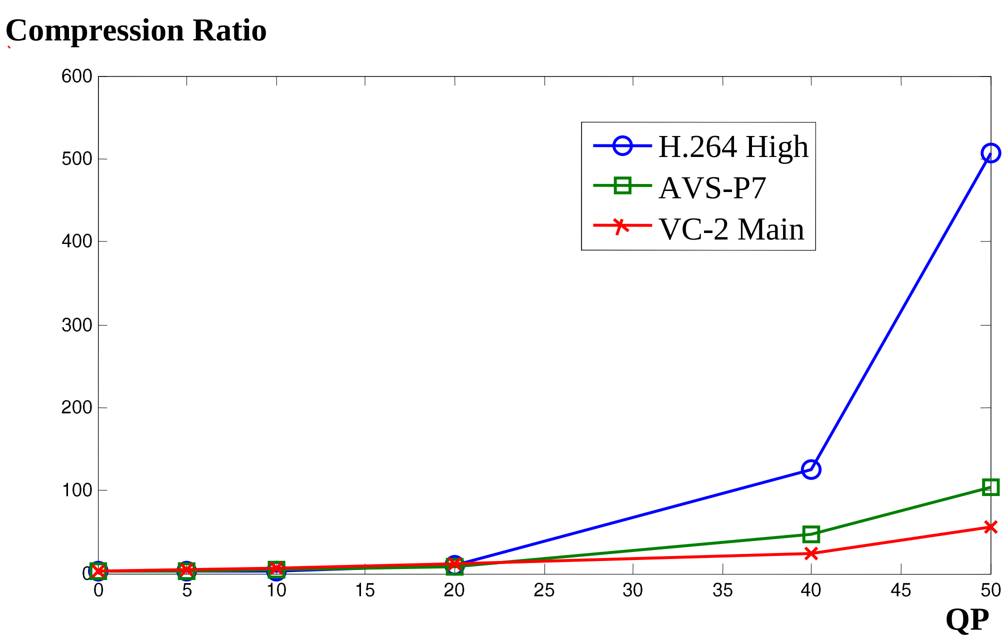


Figure 4.13 Compression Ratio of H.264 High, AVS-P7 and Dirac Pro/VC-2 Main for "Bus" sequence

Figure 4.13 shows Compression Ratio vs QP (Quantization Parameter) curves for CIF (Bus) sequence. It is observed for CIF (Bus) sequence has a same trend as QCIF sequence. As compared to QCIF sequence, CIF have a less compression ratio because of the fact that Bus CIF has lots of motion as compared to Akiyo QCIF.

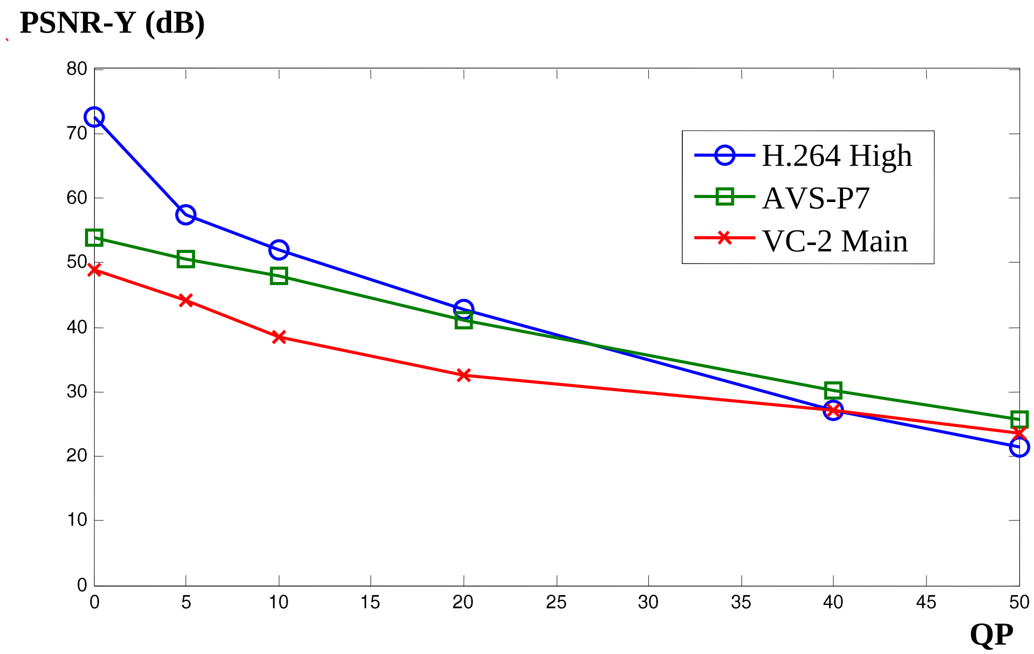


Figure 4.14 PSNR-Y of H.264 High, AVS-P7 and Dirac Pro/VC-2 Main for "Bus" sequence

Figure 4.14 shows PSNR-Y (dB) vs QP (Quantization Parameter) curves for CIF sequence. It is evident that at higher QP, H.264 performance decreases as compared to other two, the quality depreciates, i.e., distortion occurs in case of CIF sequence too. Around QP of 25-35, a suitable compression is achieved and also we get a clear winner in terms of a good PSNR-Y, i.e., AVS-P7 which performs slightly better than H.264 and much better than VC-2 in that range.

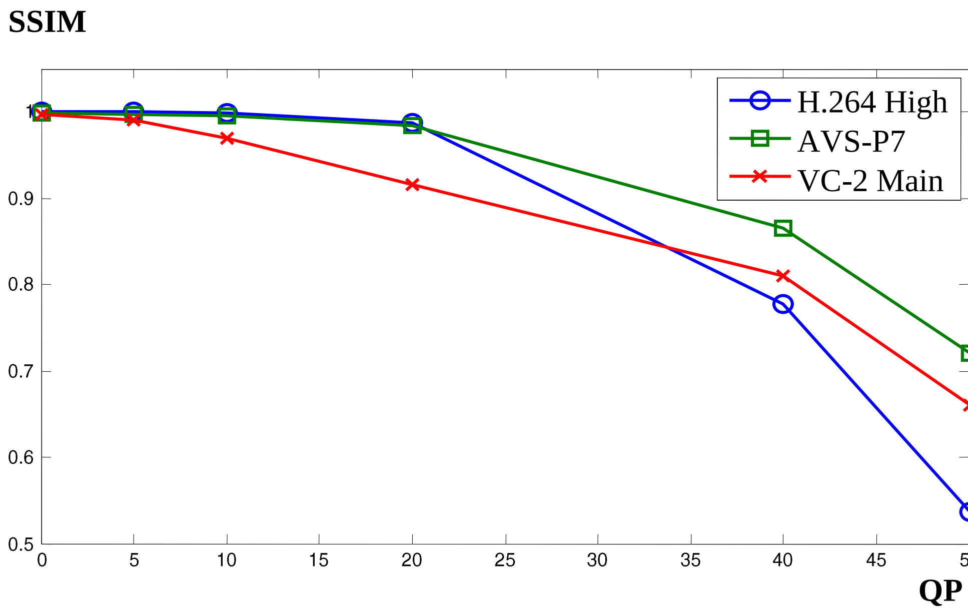
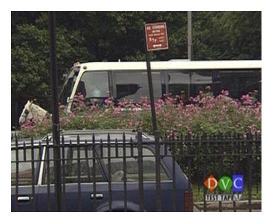


Figure 4.15 SSIM of H.264, AVS-P7 and Dirac Pro/VC-2 for "Bus" sequence

Figure 4.15 shows the SSIM vs QP (Quantization Parameter) curves for CIF (Bus) sequence. Same trend is followed as QCIF. Figure 4.16 shows the reconstructed video sequences of CIF at QP=0 and QP=50. QP=0 has no artifacts for all the three codecs but at QP=50, AVS-P7 > VC-2 > H.264, in terms of quality.

(a) H.264 (QP=0)

(b) AVS-P7 (QP=0)



(c) VC-2 (QP=0)

(d) H.264 (QP=50)

(e) AVS-P7 (QP=50)

(f) VC-2 (QP=50)

Figure 4.16 Reconstructed sequences for CIF Sequence

4.4 Impact of QP-High Bitrate Region

High bitrate region corresponds to a region where we compare only H.264 and VC-2 which encodes HD-Video (Stockholm) as AVS P7 cannot encode it. Encoding for HD video in JM is done according to Table 3-8 from Chapter 3, where as for VC-2, Table 3-5 is followed. H.264 High and VC-2 Main are used as profiles for this section.

4.4.1 Stockholm HD Sequence

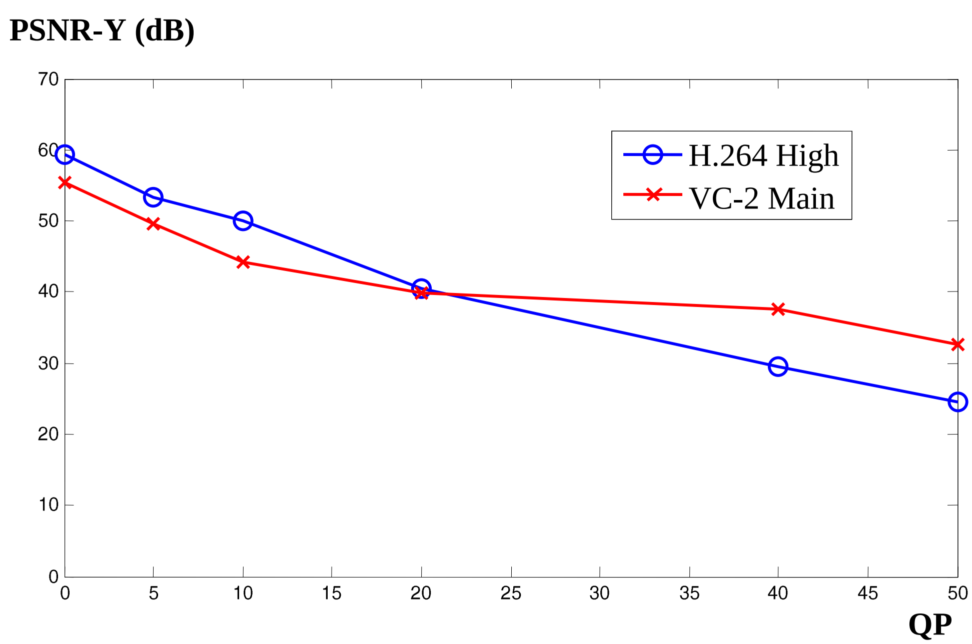


Figure 4.17 PSNR-Y of H.264 High and Dirac Pro/VC-2 Main for "Stockholm" sequence

Figure 4.17 shows the PSNR-Y (dB) curve for the HD Sequence (Stockholm). Only the first 50/239 frames are encoded. It is obvious that PSNR-Y (dB) decreases as QP increases. H.264 has considerable lead until QP=20, but VC-2 performance surpasses H.264 after that until QP=50. At QP=50, the quality starts depreciating.

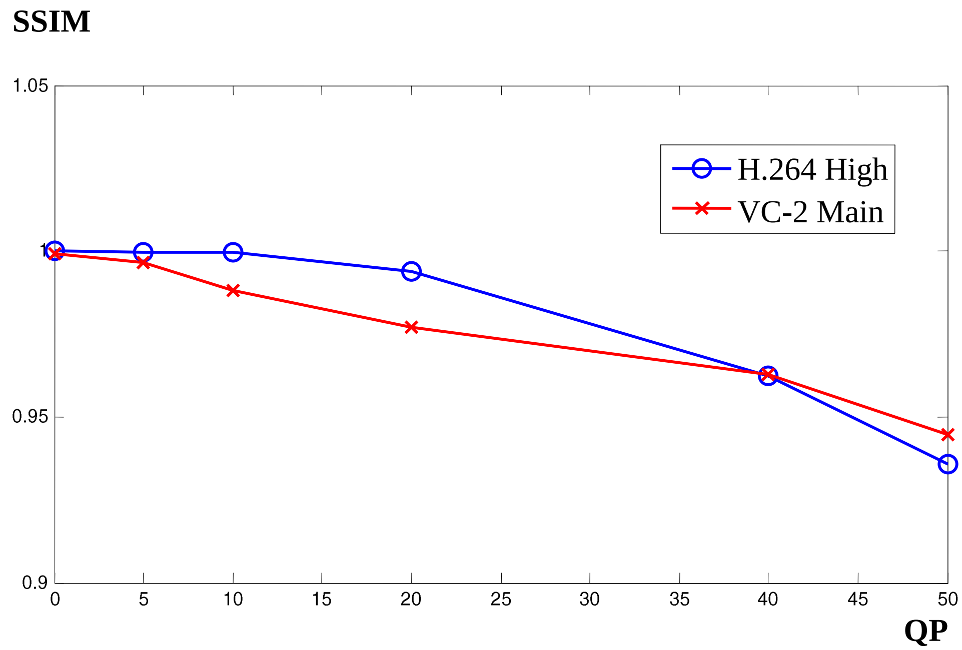


Figure 4.18 SSIM of H.264 High and Dirac Pro/VC-2 Main for "Stockholm" sequence

Figure 4.18 shows the SSIM curve for the HD sequence (Stockholm). The trend is similar to PSNR-Y (dB) curve but the VC-2 codec overtakes H.264 at QP=40. Figure 4.19 shows the reconstructed video sequence of HD at QP=0 and QP=50. QP=0 has no artifacts for all the three codecs but at QP=50, VC-2 > H.264, in terms of quality.

(a) H.264 (QP=0)

(b) VC-2 (QP=0)

(c) H.264 (QP=50)

(d) VC-2 (QP=50)

Figure 4.19 Reconstructed sequences for HD sequence

4.5 Low Bitrate Region – All Profiles

Setup for encoding the QCIF, CIF and SD in this test is followed from Section 3.1.3. H.264 (High, Baseline, Extended, Main), VC-2 (Main, Simple) and AVS-P7 are used as profiles for this section. For Section 4.5.1 and Section 4.5.2, it seems that H.264 High overlaps with H.264 Main and Extended because of a slight difference in between them.

4.5.1 Miss-America QCIF Sequence

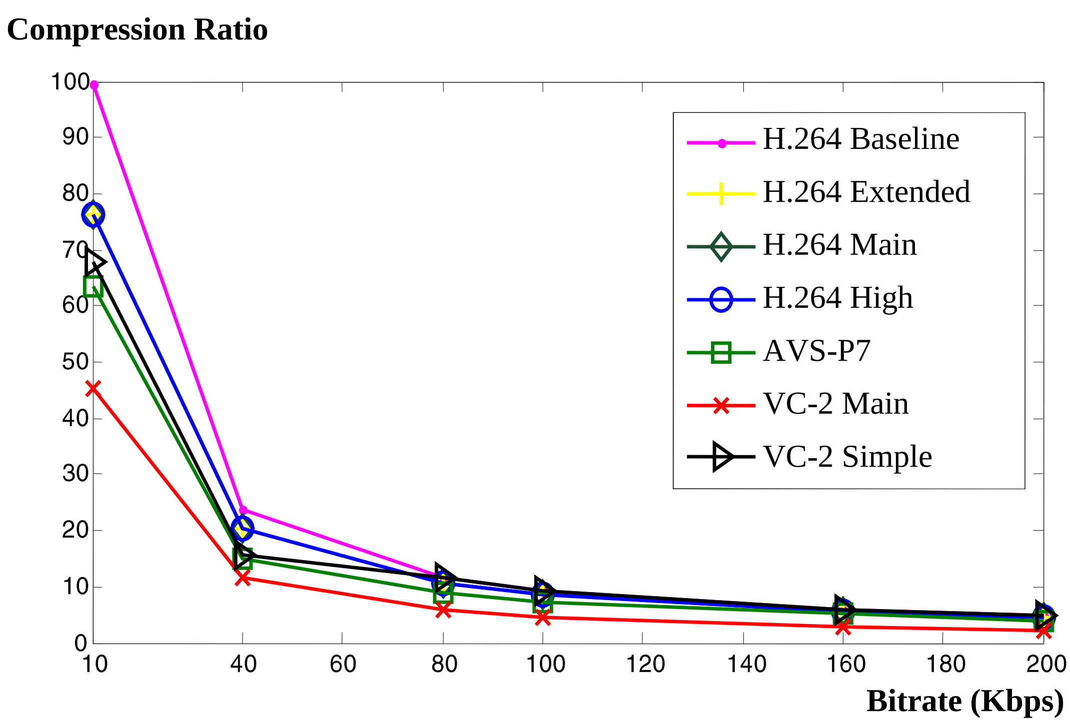


Figure 4.20 Compression Ratio of H.264, AVS-P7 and Dirac Pro/VC-2 Profiles for "Miss-America" sequence

Figure 4.20 shows the Compression Ratio vs Bitrate (Kbps) curves for QCIF (Miss-America) comparing various profiles of H.264, VC-2 with AVS-P7. For QCIF, compression is best for H.264 Baseline and worst for VC-2 Main for bitrates until 80 Kbps and uptill 200 Kbps, all the profiles saturate at a similar value except VC-2 Main.

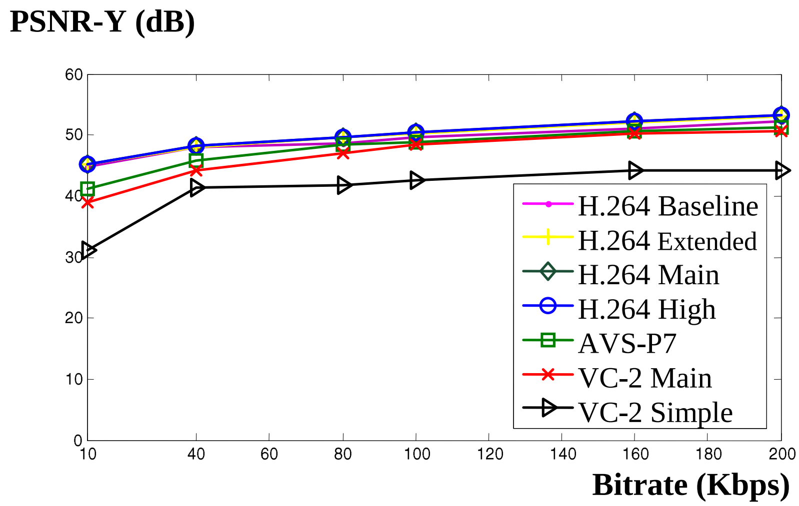


Figure 4.21 PSNR-Y of H.264, AVS-P7 and Dirac Pro/VC-2 Profiles for "Miss-America" sequence

Figure 4.21 shows the PSNR-Y (dB) vs Bitrate (Kbps) curves for QCIF (Miss-America) to compare various profiles of H.264, VC-2 with AVS-P7. For QCIF and CIF sequences, PSNR-Y increases at lower bitrates, but then saturates around 200 Kbps. It is clearly seen the order of performance is H.264 High > Main > Extended > Baseline > AVS-P7~VC-2 Main > VC-2 Simple. For QCIF sequence, VC-2 Simple has the lowest performance.

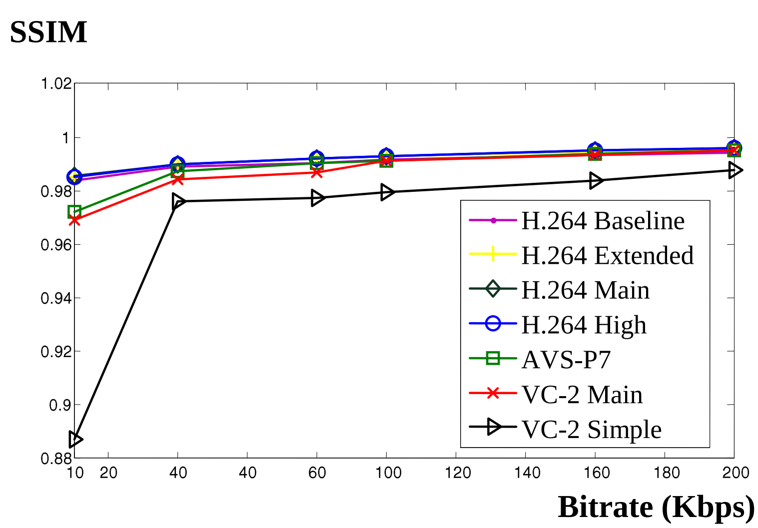


Figure 4.22 SSIM of H.264, AVS-P7 and Dirac Pro/VC-2 Profiles for "Miss-America" sequence

Figure 4.22 shows the SSIM vs Bitrate (Kbps) curves for QCIF (Miss-America) to compare various profiles of H.264, VC-2 with AVS-P7. SSIM follows PSNR-Y (dB) curve.

4.5.2 Stefan CIF Sequence

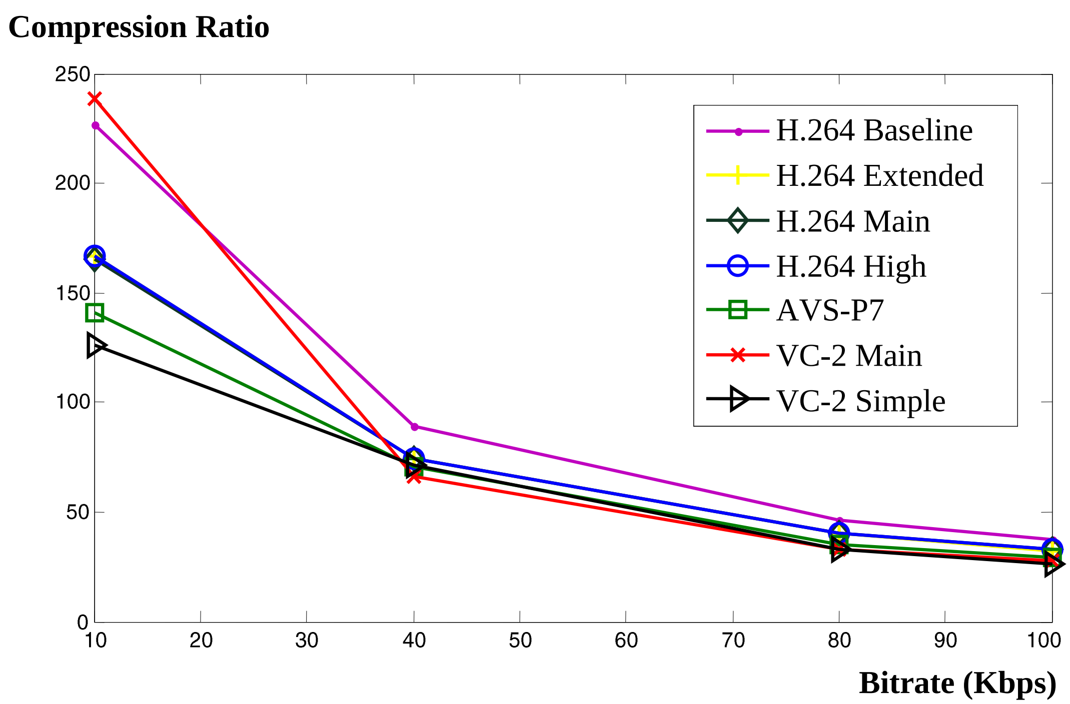


Figure 4.23 Compression Ratio of H.264, AVS-P7 and Dirac Pro/VC-2 Profiles for "Stefan" sequence

Figure 4.23 shows the Compression Ratio vs Bitrate (Kbps) curves for CIF (Stefan) to compare various profiles of H.264, VC-2 with AVS-P7. For CIF, the trend is the same except that VC-2 Main is comparable with other profiles. Compression is acceptable in the range of 40-80 Kbps, provided if it provides good PSNR-Y as shown in Figure 4.24. H.264 Baseline performs best compression among all.

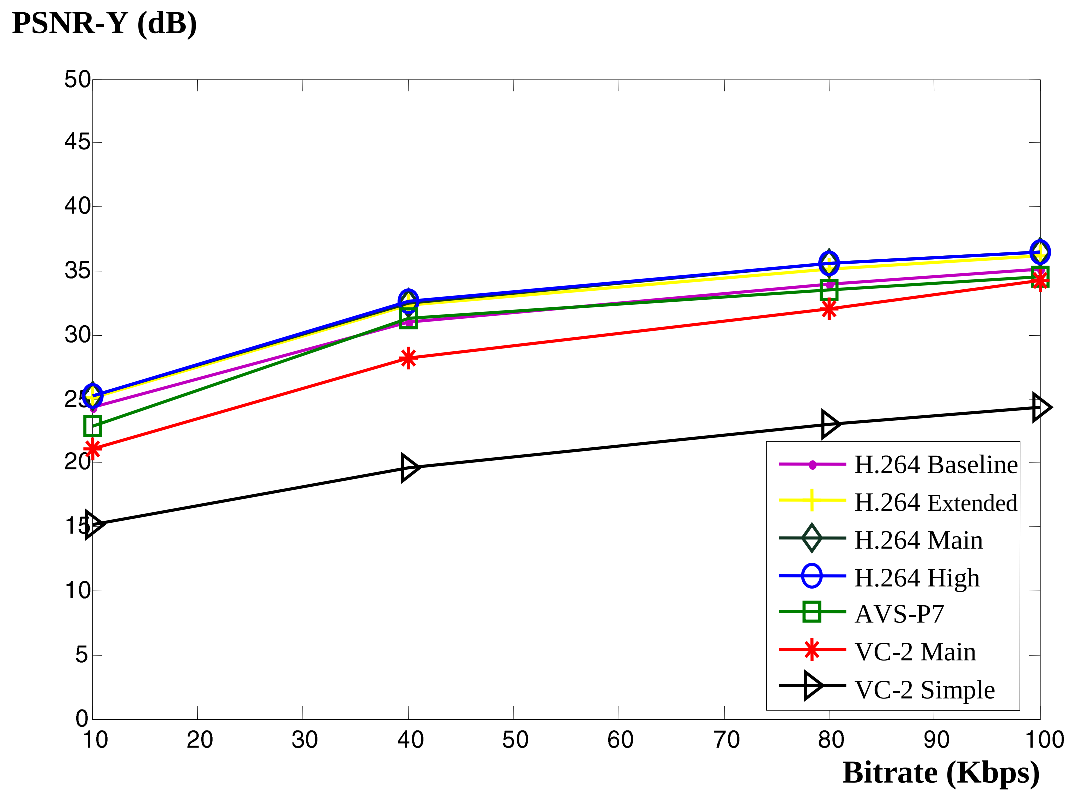


Figure 4.24 PSNR-Y of H.264, AVS-P7 and Dirac Pro/VC-2 Profiles for "Stefan" sequence

Figure 4.24 shows the PSNR-Y (dB) vs Bitrate (Kbps) curves for CIF (Stefan) to compare various profiles of H.264, VC-2 with AVS-P7. CIF sequence follows the same trend as QCIF sequence, i.e., H.264 High > Main > Extended > Baseline > AVS-P7~VC-2 Main > VC-2 Simple. The clear winner in this case is H.264 High. Around 60 Kbps, an acceptable PSNR-Y and compression is achieved. So, we get a range which is suitable for video coding for CIF sequence and can be used for producing good quality reconstruction at the receiver end and can also be transmitted through a low bandwidth channel whenever required. AVS-P7 performs moderate but equals to H.264 Main. VC-2 Main only comes closer to AVS-P7 and H.264 profiles around 100 Kbps and compression is also comparatively low. VC-2 Simple performs the lowest.

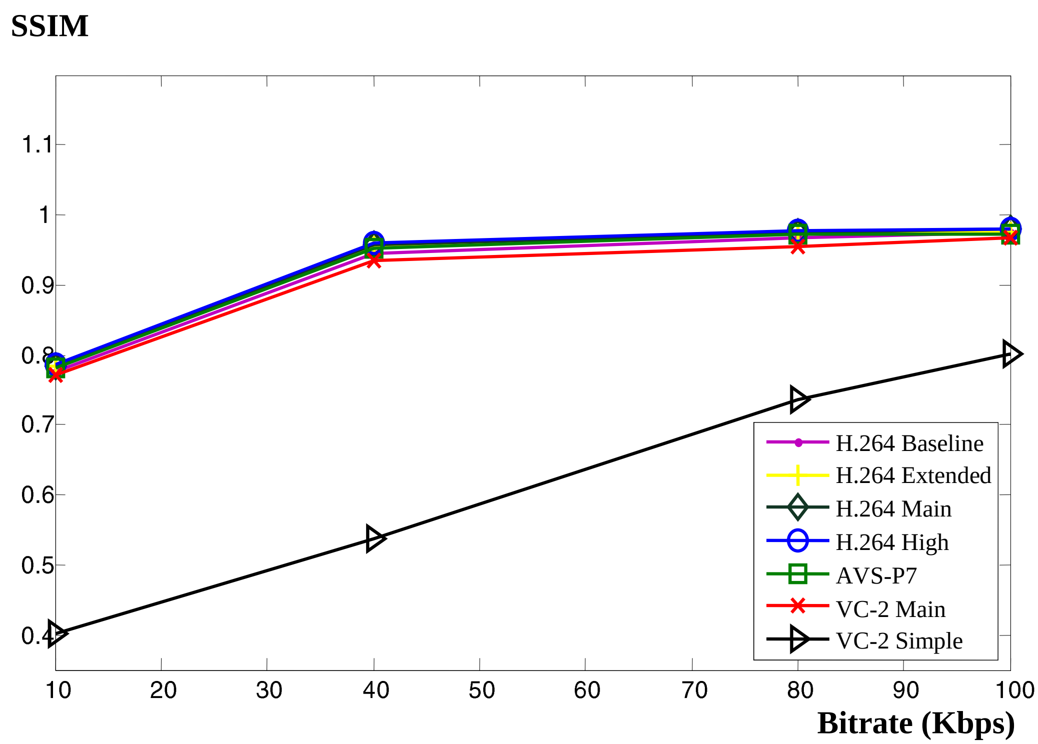


Figure 4.25 SSIM of H.264, AVS-P7 and Dirac Pro/VC-2 Profiles for "Stefan" sequence

Figure 4.25 shows the SSIM vs Bitrate (Kbps) curves for CIF (Stefan) to compare various profiles of H.264, VC-2 with AVS-P7. SSIM follows PSNR-Y (dB) curve for CIF concluding VC-2 Simple performing the lowest.

4.6 High Bitrate Region – All Profiles

Encoding for HD video in JM is done according to Table 3-3 from Chapter 3, where as for VC-2, Table 3-5 is followed. H.264 (High, Baseline, Extended, Main), VC-2 (Main, Simple) and AVS-P7 are used as profiles for this section. Analysis has been done in terms of PSNR-Y and SSIM. Compression ratio analysis has not been done for this since 50/239 frames were encoded for this sequence, because of the high encoding time for HD video in PC workstation. For Section 4.6.1, it seems H.264 High overlaps with H.264 Main and VC-2 Simple, because of a very slight difference in between them.

4.6.1 Stockholm HD Sequence

PSNR-Y (dB)

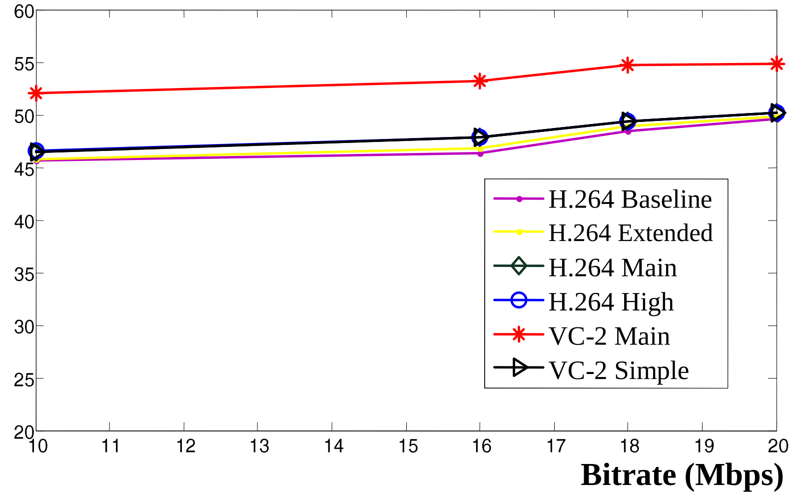


Figure 4.26 PSNR-Y of H.264 and Dirac Pro/VC-2 Profiles for “Stockholm” sequence

Figure 4.26 shows the PSNR-Y (dB) curve for HD Sequence (Stockholm). Only the first 50/239 frames are encoded, because the encoding time in PC workstation is too high. For the first 50 frames, VC-2 Main surpasses VC-2 Simple and all the profiles of H.264. Clear winner in this case is VC-2 Main performing better than H.264 Profiles.

SSIM

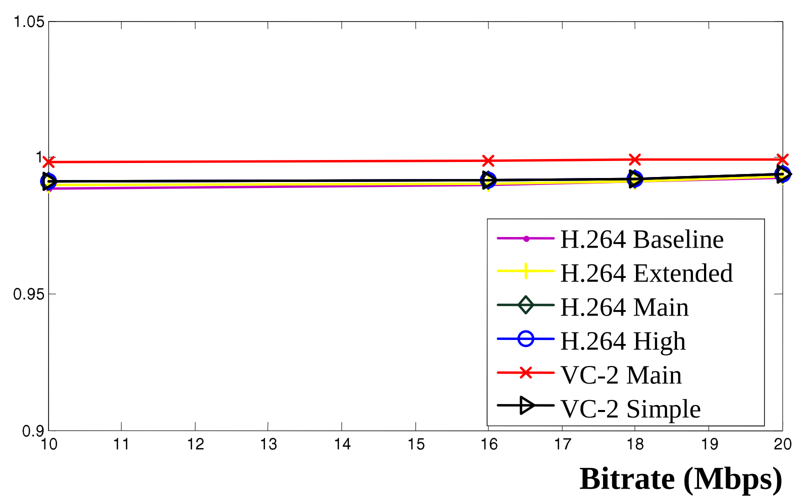


Figure 4.27 SSIM of H.264 and Dirac Pro/VC-2 for “Stockholm” sequence

Figure 4.27 shows the SSIM curve for HD Sequence (Stockholm). As PSNR-Y (dB), the same goes for SSIM.

4.7 Summary

This section summarizes the performance of different standards with respect to QP (Table 4-1) and Bitrate (Table 4-2) and covers QCIF, CIF, SD and HD sequences. Here, B – Baseline, H – High, M – Main, S – Simple, E – Extended.

For example in Table 4-1, Compression Ratio (CR) performance for QCIF at QP, i.e., Quantization Parameter = 20 is (H.264 H, VC-2 M), that means H.264 High > VC-2 Main. With respect to PSNR-Y and SSIM, the order remains the same. Table 4-2 is the performance with respect to bitrate and concludes the best, moderate and lowest codec. CR (H.264 B, H.264 H, VC-2 M) at bitrate of 40 Kbps shows that the Compression Ratio of H.264 Baseline > H.264 High > VC-2 Main at 40 Kbps.

Table 4-1 Order of performance of H.264 AVC and Dirac Pro/VC-2 by varying QP

Sequence	**QP**	**CR**	**PSNR-Y**	**SSIM**
QCIF (176x144)	20	H.264 H VC-2 M	H.264 H VC-2 M	H.264 H VC-2 M
	45	H.264 H VC-2 M	VC-2 M H.264 H	VC-2 M H.264 H
CIF (352x288)	20	H.264 H VC-2 M	H.264 H VC-2 M	H.264 H VC-2 M
	45	H.264 H VC-2 M	VC-2 M H.264 H	VC-2 M H.264 H
SD (704x576)	20	H.264 H VC-2 M	H.264 H VC-2 M	H.264 H VC-2 M
	45	H.264 H VC-2 M	VC-2 M H.264 H	VC-2 M H.264 H
HD (1280x720)	20	First 50 frames were encoded	H.264 H VC-2 M	H.264 H VC-2 M
	45		VC-2 M H.264 H	VC-2 M H.264 H

Table 4-2 Order of Performance of H.264 AVC and Dirac Pro/VC-2 Profiles by varying Bitrate

Order of Performance (Best, Moderate, and Lowest)				
Sequence	**Bitrate**	**CR**	**PSNR-Y**	**SSIM**
QCIF (176x144)	40Kbps	H.264 B H.264 H VC-2 M	H.264 H H.264 E VC-2 S	H.264 H H.264 E VC-2 S
	200Kbps	H.264 B H.264 H VC-2 M	H.264 H H.264 E VC-2 S	H.264 H H.264 E VC-2 S
CIF (352x288)	40Kbps	H.264 B H.264 H VC-2 M	H.264 H H.264 E VC-2 S	H.264 H H.264 E VC-2 S
	200Kbps	H.264 B H.264 H VC-2 M	H.264 H H.264 E VC-2 S	H.264 H H.264 E VC-2 S
SD (704x576)	2Mbps	H.264 B H.264 M VC-2 M	H.264 H H.264 E VC-2 S	H.264 H H.264 E VC-2 S
	8Mbps	H.264 B H.264 M VC-2 M	H.264 H H.264 E VC-2 S	H.264 H H.264 E VC-2 S
HD (1280x720)	10Mbps	First 50 frames were encoded	VC-2 M H.264 E H264 B	VC-2 M H.264 E H264B
	20Mbps		VC-2 M H.264 E H264 B	VC-2 M H.264 E H264 B

CHAPTER 5. CONCLUSION AND FUTURE WORK

5.1 Conclusion

Performance analysis on the three mainstreams video coding Standards, H.264 AVC and Dirac Pro/VC-2, AVS-P7 are presented. For the sequences QCIF, CIF and SD which are classified in low bitrate region with the increment in bitrate, the three standards converge to approximate similar value in terms of CR and SSIM but the H.264 High performs slightly better in terms of PSNR-Y. For the HD sequence which is classified under high bitrate region and differentiates VC-2 and H.264, by encoding the first 50 frames, VC-2 Main performs better than H.264 High.

For the variable QP, it is evident that the H.264 AVC performs better in terms of CR with respect to QCIF, CIF and SD (low bitrate region) than VC-2 and AVS-P7. However, AVS-P7 surpasses VC-2 and H.264 AVC at higher QP for QCIF, CIF and SD in terms of PSNR-Y and SSIM. For HD (high bitrate region) sequence, VC-2 takes a lead over H.264 at higher QP. But, it is evident from the output that encoding a video at higher QP would probably introduce artifacts and distort the overall quality. On an average, for sequences encoded in low bitrate region, at QP of 20-30, greater quality is achieved. Let us take QP=20, compression of all the three are the same and in terms of PSNR-Y and SSIM, H.264 is better than AVS-P7 and VC-2. But around QP=30, the compression is really high for H.264 and in terms of PSNR-Y and SSIM, AVS-P7>H.264>VC-2. For HD, around QP of 20-30, at QP=20, H.264 performs better but after QP=30, VC-2 takes the lead.

Profile comparison also shows that except for the VC-2 Simple, all the other codec profiles perform similar with respect to PSNR-Y and SSIM for QCIF, CIF and SD (low bitrate region). For HD (high bitrate region), VC-2 Main performs better than H.264 profiles. Considering the fact H.264 uses much complex hardware than the other and not being royalty free, with just slight difference in the quality, Dirac Pro/VC-2 and AVS-P7 can be proven as a better option for some applications.

In overall Dirac Pro is very promising. According to BBC R&D, Dirac Pro is a royalty free technology that anyone can use. VC-2 provides efficient compression but is simple and cost effective to implement in hardware and software for a wide range of applications. Compression parameters can be chosen to optimize VC-2 for different applications in terms of factors such as latency, compression performance, and complexity (e.g. ease of implementation and cost). Its applications are such as web streaming and IP TV and desktop production. It can be used for the same applications as AVC and for other applications as well.

Dirac Pro/VC-2 performance in terms of varying bitrates is approximately similar to H.264 and AVS-P7 in terms of compression ratio, but lags behind H.264 in terms of PSNR-Y and SSIM for QCIF, CIF and SD Media. For HD (first 50 frames), VC-2 surpasses H.264 AVC (Bitrates from 10-20 Mbps and QP=20) in terms of PSNR-Y and SSIM. For HD productions and applications such as video streaming, Dirac Pro/VC-2 holds good when encoded under the target Bitrate mentioned and QP range.

On the other hand, AVS-M standard can cover a broad range of applications including mobile multimedia broadcasting, IP multimedia subsystem (IMS), multimedia mailing, multimedia services over packet networks, video conferencing, video phone, and video surveillance. It is evident that AVS-P7 handles QCIF, CIF and SD media at QP=30 better than H.264 and VC-2. Therefore AVS-P7 can have a really effective encoding and also with respect to Bitrates it performs considerably the same as other two standards. So at low bitrates between 100-200 Kbps for QCIF and CIF and around 4 Mbps for SD, AVS-M is also a good choice for an alternative codec.

It is suggested that, despite its simple toolset, Dirac Pro and AVS-P7 is very comparable to other state-of-the-art codec such as H.264 AVC. However the question remains whether the enormous cost in royalty fees justifies the additional increase in quality.

Profile comparison proves the fact that H.264 High is better when one look into profiles of H.264. VC-2 Main is better than VC-2 Simple in terms of PSNR-Y and SSIM. So when one wants an alternative codec to H.264 High, VC-2 Main and AVS-M (P7) can always be used.

5.2 Future Work

Comparative analysis can be further extended to AVS CHINA-P2 which is in parallel with H.264 and Dirac Pro/VC-2 for encoding HD Video. A Combined DCT Architecture can also be implemented in hardware. One of the basic building blocks in any video codec is the Transform Unit. From Table 2-3, all of the three have one thing in common, the 4*4 Transform. Shared architecture for the transform unit of all of the three codecs can be proposed to be implemented on an FPGA for trans-coding applications which contributes to a much lower hardware and reduces the complexity in the quantizer and dequantizer.

REFERENCES

[1] http://desktopvideo.about.com/od/glossary/g/vidcompression.htm

[2] H.264/MPEG-4 AVC: Wikipedia, accessed on Nov. 23, 2010.

[3] AVS: www.avs.org.cn, accessed on Nov. 21, 2010.

[4] Dirac Pro "Light Compression" by Tim Borer www.tech.ebu.ch/docs/events/wbuisog09/presentations/ebu_wbuis0g09_borer.pdf

[5] K. R. Rao, Do Nyeon Kim, "Current Video Coding Standards: H.264/AVC, Dirac, AVS China and VC-1," 42nd South Eastern Symposium on System Theory, University of Texas at Tyler, TX, USA, March 7-9, 2010.

[6] Wen Gao, Siwei Ma, Li Zhang, Li Su, and Debin Zhao, AVS Video Coding Standard, C.W. Chen et al. (Eds.): Intel.Multimedia Communication: Tech. and Applied, Springer- Verlag Berlin Heidelberg, pp. 125–166, 2010.

[7] Aruna Ravi, M.Sc. Thesis, "Performance Analysis and Comparison of Dirac Video Codec with H.264 / MPEG-4 PART 10 AVC," University of Texas-Arlington, August 2009.

[8] PSNR: http://en.wikipedia.org/wiki/Peak_signal-to-noise_ratio.

[9] SSIM: https://ece.uwaterloo.ca/~z70wang/research/ssim/

[10] S. K. Kwon, A. Tamhankar and K. R .Rao, "Overview of H.264 / MPEG-4 Part 10" J. Visual Communication and Image Representation, Vol 17, pp. 186-216, April 2006.

[11] H.264: http://www.drtonygeorge.com/, accessed on Nov 20, 2010.

[12] Dirac Pro: www.bbc.co.uk/rd/projects/dirac/diracpro.shtml, accessing on Nov. 23, 2010.

[13] http://dirac.sourceforge.net/documentation/algorithm/algorithm/overall_arch.htm

[14] Wen Gao, Siwei Ma, Li Zhang, Li Su, and Debin Zhao, AVS Video Coding Standard, C.W. Chen et al. (Eds.): Intel. Multimedia Communication: Tech. and Appli., Springer- Verlag Berlin Heidelberg, pp. 125–166, 2010.

[15] Liang F., Mobile Multimedia Broadcasting Standards: Technology and Practice, 485 DOI: 10.1007/978-0-387-78263-8 17, Springer Science+Business Media, LLC 2009.

[16] JM: http://iphome.hhi.de/suehring/tml/, accessed on Oct. 29, 2010.

[17] Schroedinger: http://diracvideo.org/, accessed on Oct. 29 2010.

[18] http://ee.uta.edu/Dip/Courses/EE5359/index.html, accessed on Nov. 17, 2010.

[19] http://www.eng.tau.ac.il/~yaro/adiplab/Lab4_Transform Coding.pdf, accessed on Nov. 25, 2010.

[20] http://www.vcodex.com/h264transform4x4.html, accessed on Nov. 23, 2010.

[21] http://dirac.sourceforge.net/documentation/code/programmers_guide/toc.htm, accessed on Nov 25, 2010.

[22] QCIF and CIF: http://trace.eas.asu.edu/yuv/, accessed on Oct. 29, 2010.

[23] SD and HD: www.nsl.cs.sfu.ca/wiki/index.php/Video Library and Tools, accessed on Oct. 29, 2010.