JCSE International Journal of Computer Sciences and Engineering Open Access
Review Paper Volume-2, Issue-4 E-ISSN: 2347-2693

A Review on H.264/AVC Mpeg 4 Part-10 Compression Methods

Urvashi[1*] and Kiran Jain[2]

[1*,2]*Dept. of Computer Science*, Doon Valley College of Engg. & Tech., Karnal, India

www.ijcseonline.org

Received: 19/03/2014 Revised: 26/03/2014 Accepted: 21/04/2014 Published: 30/04/2014

Abstract: H.264/AVC is newest video coding standard of the ITU-T Video Coding Experts Group and the ISO/IEC Moving Picture Experts Group. The main goals of the H.264/AVC standardization effort have been enhanced compression performance and provision of a "network-friendly" video representation addressing "conversational" (video telephony) and "non-conversational" (storage, broadcast, or streaming) applications. H.264/AVC has achieved a significant improvement in rate-distortion efficiency relative to existing standards. This article provides an overview of the technical features of H.264/AVC, describes profiles and applications for the standard, and outlines the history of the standardization process.

***Keywords**:* Video Compression Systems, ITUT H.264, Compression Standards

I. INTRODUCTION

The digital video compression technology has been gaining popularity for many years. Today, when people enjoy HDTV (high definition television), movie broadcasting through Internet or the digital music such as MP3, the convenience that the digital video industry brings to us cannot be forgotten. All of these should attribute to the advances in compression technology, enhancement on mass storage media or streaming video/audio services.

1.1 H.261

H.261 is first developed by ITU-T in 1990. It is a video compression standard, which targets on low bit- rate real time applications (down to 64 kbit/s), such as visual telephone service. The basic idea of video coding is based on DCT, VLC entropy coding and simple motion estimation technique for reducing the redundancy of the video information.

1.1 MPEG -1

The MPEG-1 standard, published in 1992, was designed to produce reasonable quality images and audio at low bit rates. MPEG-1 provides the resolution of 352x240 (SIF) for NTSC or 352x288 for PAL at 1.5 Mb/s. The target applications are focused on the CD-ROM, video-CD, and stream media applications like video over digital telephone networks, video on demand (VOD) etc. The picture quality level almost equals to VHS tape. MPEG-1 can also be encoded at bit rates as high as 4-5Mbits/sec. MPEG-1 specified the compression of audio signals, called layer-1,-2,-3. Layer-3 is now very popular in the digital music distribution over Internet known as MP3.

1.2 H.262 and MPEG-2

MPEG–2 standard was established by ISO/IEC in 1994. The purpose of this standard is to produce enhanced data rate and better video quality compared to MPEG–1. The coding technique of MPEG-2 is the same as MPEG-1 but with a higher picture resolution of 720x486.The unique feature of MPEG-2 is the layered structure, which supports a scalable video system. MPEG-2 is compatible with MPEG-1, that means a MPEG-2 player can play back MPEG-1 video without any modification. This standard is also adopted by ITU-T referred to as H.261.

1.3 MPEG-4

MPEG-4 (ISO/IEC 14496) became the international standard in 1999. The basic coding theory of MPEG-4 still remains the same as previous MPEG standards but more networks oriented. It is mostly used for broadcast, interactive and conversational environments. MPEG-4 introduced 'objects' concept: A video object in a scene is an entity that a user is allowed to access (seek, browse) and manipulate (cut and paste). It serves from (2 kbit/s for speech, 5 kbit/s for video) to (5 Mbit/s for transparent quality video and 64 kbit/s per channel for CD quality audio).

MPEG-4 part-10/ H.264

ITU-T Video Coding Experts Group (VCEG) and the ISO/IEC MPEG jointly develop the newest standard, H.264/AVC (also known as MPEG-4 part 10). The motivation of this standard comes from the growing multimedia services and the attractiveness of HDTV, which need more proficient coding method. At the same time, various transmission media especially for those low speed media (Cable Modem, xDSL or UMTS) also called for the significant enhancement of coding efficiency.
By introducing some unique techniques, H.264/AVC aims to increase compression rate significantly (save up to 50% bit rate as compared to MPEG-2 picture quality) while transmitting high quality image at both high and low bit rates. The standard can increase resilience to errors by supporting flexibility in coding as well as organization of coded data.

Architecture of H.264

H.264/AVC is the latest international standard for video coding, issued in May 2003. It was jointly developed by the ITU-T Video Coding Experts Group (VCEG) together with the ISO/IEC Moving Picture Experts Group (MPEG).

Corresponding Author: S.G. Santhi
Dept. of CS, Doon Valley College of Engg. & Tech., Karnal, India

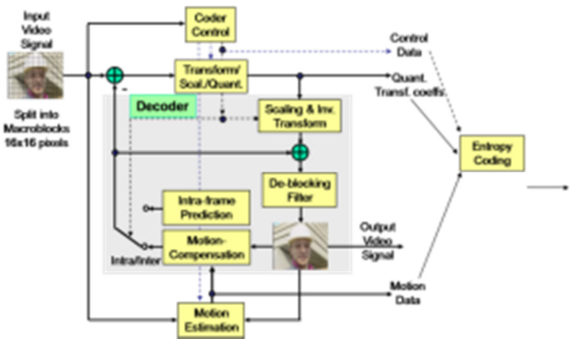

Figure 1: The ITUT H.264 MPEG-10 AVC Video Codec Overview

The official name is Advanced Video Coding (AVC), also known as H.264 or MPEG 4 Part 10. The standard defines the video bit-stream and decoding method, allowing design flexibility for encoding process. Figure 1 briefly summarizes the ITUT H.264 MPEG-10 AVC series video coding standards

Compared to the other standards, H.264/AVC contains a number of new features, which not only offers lower bit rate and more efficient compression, but also provide more flexibility for application to a wide variety of network environments. As shown in Figure 2, H.264 consists of two layers, namely Video Coding Layer (VCL), and Network Abstraction Layer (NAL).

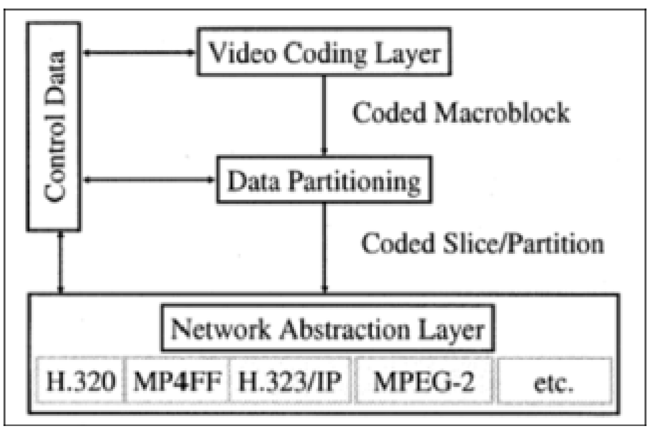

Figure 2: Two layers in H.264/AVC.

Like the other video coding standards, H.264/AVC incorporates different profiles and levels. There are up to 16 profiles and 16 levels in the current version. Three most commonly used profiles are baseline profile (BP), main profile (MP), and extended profile (EP).

1.4 Applications of H.264

The H.264 was designed to be flexible video format and has a very broad application range including:

1. Broadcast over cable modem, satellite, cable, terrestrial, DSL, etc.
2. Interactive or serial storage on DVD, optical and magnetic devices, etc.
3. Conversational services over Ethernet, LAN, ISDN, DSL, wireless and modems, mobile networks, etc. or mixtures of these.
4. Video-on-demand or multimedia streaming services over ISDN, cable modem, DSL, LAN, wireless networks, etc.
5. Multimedia messaging services (MMS) over Ethernet, ISDN, DSL, mobile networks, LAN, and wireless, etc.
6. Low bit-rate Internet streaming applications.
7. HDTV broadcast and Digital Cinema applications.
8. Web software Embedding.
9. Mobile TV standardization.
10. Video conferencing products.

1. Advantages of H.264

With H.264 Codec Longer record times is Possible: H.264 provides magnificent compression, in almost cases more than doubling record times over previously popular compression methods.

Improved Quality and speed: Previous compression methods have always allowed excellent picture quality at real-time frame rates but they take up valuable hard drive space.

H.264 provides improved remote monitoring: H.264 not only Save hard drive space but also provide the ability to bring together high quality and low memory sizes allows for presentations of video when transmitted.

Error Robustness: The H.264 decoder solution has built with robust error handling. The whole stream is divided into independent NAL unit, if some errors are detected then the decoder stops decoding that particular NAL unit packet and continues to decode the next good NAL unit packet.

Design Flexibility & Modularity: H.264 decoder solution is very flexible and modular to suit the requirements of wide-array of broadcast / professional video & surveillance applications. H.264 decoder solution can be customized to accomplish decoding of multiple streams at once.

II. RELATED WORK

Wang et al[1] describes the structural similarity (SSIM) index has been found to be a good indicator of perceived image quality. In this paper, they propose a rate-SSIM optimization scheme for mode selection in H.264/AVC video coding. The proposed method is fully standard-compatible. Experimental results demonstrate that, compared with conventional rate distortion optimization coding schemes, the proposed scheme can achieve better rate-SSIM performance and provide better visual quality.

Sequence		$\Delta SSIM$	ΔR^*	$\Delta PSNR$
Silent (CIF)	IPP..	0.0115	-14.62%	-0.14dB
	IBP..	0.0064	-8.07%	-0.25dB
Flower (CIF)	IPP..	0.0076	-14.34%	-0.66dB
	IBP..	0.0034	-6.73%	-0.55dB
Bus (CIF)	IPP..	0.0136	-14.71%	-0.51dB
	IBP..	0.0081	-8.95%	-0.62dB
Salesman (QCIF)	IPP..	0.0185	-17.09%	0.08dB
	IBP..	0.0096	-8.45%	-0.15dB
Carphone (QCIF)	IPP..	0.0038	-6.89%	-0.47dB
	IBP..	0.0008	-2.11%	-0.67dB
Container (QCIF)	IPP..	0.0087	-17.23%	0.06dB
	IBP..	0.0049	-12.41%	-0.26dB
Average	IPP..	0.0106	-14.15%	-0.27dB
	IBP..	0.0055	-7.79%	-0.42dB

* ΔR in terms of SSIM.

Table:1 Performance of the proposed scheme (compared with the original RDO technique).

They propose an RDO scheme for H.264/AVC video coding, aiming for achieving the best rate-SSIM performance. The novelty of our approaches lies in the adaptive Lagrange multiplier selection methods at the frame level, where they incorporated a new RR-SSIM estimation algorithm and a source-side information combined rate model. Our experiments show that the proposed scheme offers significant rate reduction while keeping the same level of SSIM quality value.

Dembla et al[2] describes the data quantity is very large for the digital video and the memory of the storage devices and the bandwidth of the transmission channel are not infinite, so it is not practical for us to store the full digital video without processing. This paper starts with an explanation of the basic concepts of video compression algorithms and then introduces and performs video compression standards H.264 and MPEG4. In paper highly flexible approach of H.264 & MPEG4 concentrates specifically on efficient compression of video frames base on PSNR.

Table 1 PSNR ratio of Input file/output file which is decoded by H.264/AVC codec

	H.264/AVC (Kbps)		
Frame Number	Y Sample	U Sample	V Sample
1	46.394039	53.996105	53.960239
2	45.490421	52.379498	52.67915
3	45.558582	52.31337	52.632984
4	45.626572	52.256119	52.971619
5	46.016186	52.633461	53.199165
6	46.525066	53.072243	53.385113
7	46.661331	53.100918	53.474575
8	46.680725	53.23785	53.463364
9	47.327545	52.922421	53.013885
10	48.199886	53.446217	53.747978
11	48.510048	53.598457	53.995232
12	51.058014	56.124378	56.917618
13	50.453037	55.10667	55.68462
14	49.531715	53.82954	54.601696
15	48.377861	52.510666	53.460659
16	49.13126	52.998768	53.605232
17	48.955276	52.881874	53.409069
18	48.806713	52.888805	53.390995
19	48.99382	52.921059	53.249226
20	48.844971	53.223209	53.555061
21	49.071579	52.920208	53.397266
22	48.915665	52.852592	53.118
23	48.904636	52.556797	53.034824
24	50.091148	54.425903	55.320705
25	49.558479	53.583546	54.404991

Table 2 PSNR ratio of Input file/output file which is decoded by MPEG-4 codec

	MPEG-4 (Kbps)		
Frame Number	Y Sample	U Sample	V Sample
1	55.468803	65.050385	62.511284
2	54.318981	62.425312	61.731377
3	53.681393	61.127552	60.438801
4	53.474285	60.660854	60.390961
5	53.050465	59.83023	59.578423
6	52.881157	59.411217	59.330711
7	52.508533	58.9986	58.915836
8	52.44199	58.661243	58.868668
9	51.7272	57.458427	57.761833
10	51.628216	56.930038	57.109093
11	51.475655	56.659435	56.816589
12	67.18438	71.232536	74.731323
13	57.996479	63.70266	64.284622
14	55.157059	59.568172	59.762192
15	52.114658	57.021049	57.86858
16	51.876286	56.144016	56.942921
17	51.412262	55.621651	56.314453
18	51.292988	55.318043	55.840641
19	51.119892	55.21685	55.721802
20	51.405136	55.341484	55.789219
21	51.018719	55.07494	55.490685
22	50.812237	54.501984	55.064877
23	50.687057	54.43845	55.047882
24	68.042641	72.638725	74.336235
25	55.16898	60.458084	60.999794
26	53.965385	58.052902	58.75359
27	53.380753	57.124763	57.925919
28	53.349094	56.86871	57.894207
29	53.243076	56.752716	57.915142
30	50.940994	54.465359	56.243202

Table:2 Comparision of PSNR ratio of H.264 and MPEG-4

Now this table shown below, shows testing for 1 to 30 (*.yuv (4:2:0)) frames and reading about input raw file & generated raw file.

Different choices during the design of a CODEC and different strategies for coding control can lead to significant variations in compression and computational performance between CODEC implementations. However, the best performance that may be achieved by a CODEC is limited by the available coding tools.

Seo, Sangwon et al[3] describes in recent years, the increasing demands of multimedia services on the cellular networks have accelerated this trend. This paper presents a low power SIMD architecture that has been tailored for efficient implementation of H.264 encoder/decoder kernel algorithms. The proposed architecture increases the throughput of H.264 encoder/decoder kernel algorithms by a factor of 2.13 while achieving 29% of energy-delay improvement on average compared to our previous SIMD architecture, SODA.

			Area		H.264 Decoder	
	Components	Units	Area mm^2	Area %	Power mW	Power %
PE	SIMD Data Mem (32KB)	2	4.88	34.15%	5.14	6.29%
	SIMD Register File (16x1024bit)	2	1.59	11.13%	14.95	18.27%
	SIMD ALUs, Multipliers, and SSN	2	2.25	15.75%	22.43	27.41%
	SIMD Pipeline+Clock+Routing	2	0.59	4.13%	11.68	14.28%
	SIMD Buffer (128B)	2	0.41	2.87%	1.70	2.08%
	SIMD Adder Tree	2	0.09	0.63%	0.52	0.64%
	Intra-processor Interconnect	2	0.47	3.29%	4.67	5.71%
	Scalar/AGU Pipeline & Misc.	2	0.61	4.27%	6.72	8.21%
System	ARM (Cortex-M3)	1	0.6	4.20%	2.5	3.05%
	Global Scratchpad Memory (128KB)	1	1.8	12.5%	10	12.22%
	Inter-processor Bus with DMA	1	1.0	7.00%	1.5	1.83%
Total	90nm (1V @300MHz)		**14.29**	100%	**81.81**	100%
Est.	65nm (1V @ 300MHz)		**7.46**		**25.76**	
	45nm (1V @ 300MHz)		**3.89**		**20.36**	

Table: 3 Summary of area and power running H.264 CIF video at 30FPS.

The mobile multimedia processor requires high performance low-power solutions for high quality video and wireless protocols. Our results show that they can achieve 2.13x speedup and 29% energy-delay improvement for the H.264 codec over a wide-SIMD architecture, SODA.

Cheng et al[4] describes Video compression deals with compact representations of video signals for storage and transmission. It takes advantage of features of the human visual system (HVS) for a more efficient compression. In this thesis, pixels covering the smooth part of partially-textured macroblocks are denoted as “target pixels”. The main objective in this thesis is to reduce ringing artifacts by compensating target pixels for distortion. The thesis algorithm is implemented in the Rate-Distortion (RDO) mode decision part in an H.264 encoder. By using this algorithm, RDO intends to select higher bit-cost modes for partially textured macroblocks (which contain target pixels), such that the distortion for target pixels is reduced. This results in less distortion and a reduced amount of ringing artifacts.

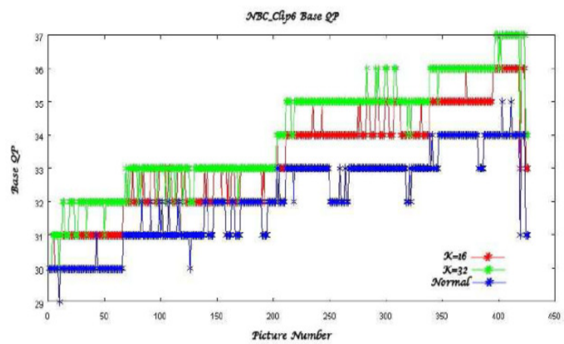

Figure:3 Base QP increases as the picture number increasing.

In this thesis project, the video CODEC system and H.264 standard has been studied, as well as the rate-distortion theory. To achieve the primary goal of reducing ringing artifacts in coded video sequences, methods were designed to pick out target pixels and compensate them for distortion.

Shafique et al[5] describes the H.264/AVC video coding standard features diverse computational hot spots that need to be accelerated to cope with the significantly increased complexity compared to previous standards. In this paper, they propose an optimized application structure (i.e. the arrangement of functional components of an application determining the data flow properties) for the H.264 encoder which is suitable for application-specific and reconfigurable hardware platforms.

For a MIPS processor they achieve an average speedup of approximately 60× for Motion Compensated Interpolation. The increases the amount of available reconfigurable hardware per Special Instruction (within a functional block) which leads to a 2.84× performance improvement of the complete encoder when compared to a Benchmark Application with standard optimizations. They evaluate our application structure by means of four different hardware platforms.

Functional component	Special instruction	Description of special instructions	Accelerating data paths
Motion estimation (ME)	SAD	Sum of absolute differences of a 16×16 macroblock	SAD_16
	SATD	Sum of absolute transformed differences of a 4×4 sub-block	QSub, Transform, Repack, SAV
Motion compensation (MC)	MC_Hz_4	Motion compensated interpolation for horizontal case for 4 pixels	PointFilter, BytePack, Clip3
Intra prediction (IPred)	IPred_HDC	16×16 intra prediction for horizontal and DC	PackLBytes, CollapseAdd
	IPred_VDC	16×16 intra prediction for vertical and DC	CollapseAdd
(Inverse) transform	(I)DCT	Residue calculation and (inverse) discrete cosine transform for 4×4 sub-block	Transform, Repack, (QSub)
	(I)HT_2×2	2×2 (inverse) Hadamard transform of Chroma DC coefficients	Transform
	(I)HT_4×4	4×4 (inverse) Hadamard transform of intra DC coefficients	Transform, Repack
Loop filter (LF)	LF_BS4	4-pixel edge filtering for in-loop de-blocking filter with boundary strength 4	Cond, LF_4

Table:4 Implemented special instructions and data paths for the major functional components of H.264 video encoder.

They have presented optimizations for the H.264 encoder application structure for reduced processing and reduced hardware pressure along with several novel data paths and the resulting. For in-loop De-blocking Filter, the optimized filtering data path reduces the number of required slices to 67.8% (i.e. 1.47× reduction, see Table 4). The Special Instruction is 120× faster than the General Purpose Processor Implementation.

Kulikov, DrDmitriy et al[6] describes the main goal of this report is the presentation of a comparative evaluation of the quality of new. The main task of the comparison is to analyze different H.264 encoders for the task of transcoding video—e.g., compressing video for personal use. Speed requirements are given for a sufficiently fast PC; fast presets are analogous to real-time encoding for a typical home-use PC.

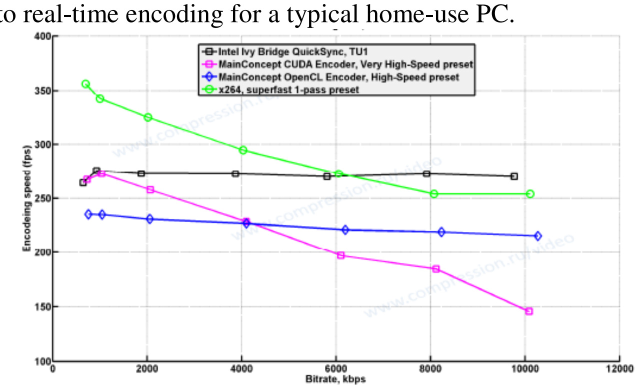

Figure: 4Encoding speed handling, Fast encoders,” “Underwater” sequence

Comparing all results from Fast encoders comparison part one could say that x264 and Intel QuickSync are best in terms of speed/quality trade-off. Main concept Open CL is third and main concept CUDA is fourth. So best of hardware encoders Intel QuickSync and best of software encoders x264 are comparable by speed/quality at very high speed encoding.

Brandão, Tomás et al[7] describes this paper describes and compares a set of no-reference quality assessment algorithms for H.264/AVC encoded video sequences. In order to obtain perceived quality scores from the estimated error, three methods are presented: i) to theyight the error estimates according to a perceptual model; ii) to linearly combine the mean squared error (MSE) estimates with additional video features; iii) to use MSE estimates as the input of a logistic function.

Metric	RMS	CC	RC	OR
Perceptual error weighting	0.52	0.91	0.92	0.14
Linear model	0.64	0.86	0.86	0.16
Logistic model	0.32	0.97	0.96	0.05
PSNR	0.83	0.75	0.76	0.28

Table:5 Evaluation of the described metrics.

Three different no-reference video quality assessment algorithms have been described and evaluated. Those algorithms share a common component on their architecture – they all use an algorithm that computes an estimation for the error due to lossy video encoding. The algorithms' performances have been evaluated using a cross-calibration procedure over 58 subjective test conditions (which are H.264 encoded versions of 12 different video sequences).

Prasantha et al[8] describes H.264/AVC (Advanced Video Coding) is the newest video coding standard of the moving video coding experts group. The paper proposes to port the H.264/AVC decoder on the various processors such as TI DSP (Digital signal processor), ARM (Advanced risk machines) and P4 (Pentium processors). The paper also proposes to analyze and compare Video Quality Metrics for different encoded video sequences.

parameters	With deblocking filter			Without deblocking filter		
	TI DSP	ARM	P4	TI DSP	ARM	P4
MSE	2.8	57.4	2.8	20.24	55.96	60.8
PSNR	43.75	30.55	43.75	35.07	30.65	30.29
SSIM	0.986	0.87	0.986	0.94	0.873	0.866
MSAD	1.03	5.065	1.03	3.136	5.0	5.14

Table:6 Video quality measures with and without deblocking filter for akiyo.

The H.264 decoder is implemented on ARM9, TMS320DM642 and Pentium 4 processor. Various parameters such as PSNR, SSIM, MSAD and MSE are calculated for the different video sequences on the three processors. From table, TI DSP performs better than the other processors for implementing H.264.

III. CONCLUSION

Although H.264/AVC is 2 -3 times more complex than MPEG-2 and the decoder is4 - 5 times more complex than the encoder, it is relatively less complex than MPEG-2 was at its outset, due to the huge progress in technology which has been made since then. Another important fact is that H.264/AVC is a public and open standard. Every manufacturer can build encoders and decoders in a competitive market. There is no dependency on proprietary formats, as on the Internet today, which is of utmost importance for the broadcast community.

IV. FUTURE SCOPE

Like the other video coding standards, H.264/AVC incorporates different profiles and levels. Profiles define sets of bit stream features a H.264 stream can use. Levels define restrictions on the video resolution, frame rate and some stuff called VBV (Video Buffer Verifier). There are up to 16 profiles and 16 levels in the current version. Three most commonly used profiles are baseline profile (BP), main profile (MP), and extended profile (EP), Two of the most commonly used profiles i.e. baseline profile (BP), main profile (MP), will be studied and SSIM (Structural Similarity) and PSNR Matrices will be analyzed for various input videos samples. In future two of the most commonly used profiles i.e. baseline profile (BP), main profile (MP), will be studied and SSIM (Structural Similarity) and PSNR Matrices will be analyzed for various input videos samples.

V. REFERENCES

[1]. Wang, Shiqi, Abdul Rehman, Zhou Wang, Sitheyi Ma, and TheynGao. "Rate-SSIM optimization for video coding." In Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on, pp. 833-836. IEEE, 2011.

[2]. Dembla, Deepak, Biren Patel, Ashish Kumar, and YogeshBhomia. "Comparison of H. 264 and MPEG-4 Codec Based on PSNR-Peak Signal to Noise Ratio Algorithm." International Journal 3, no. 3 (2013).

[3]. Seo, Sangwon, Mark Woh, S. Mahlke, T. Mudge, Sunfaram Vijay, and ChaitaliChakrabarti. "Customizing wide-SIMD architectures for H. 264." In Systems, Architectures, Modeling, and Simulation, 2009. SAMOS'09. International Symposium on, pp. 172-179. IEEE, 2009.

[4]. Cheng, Xiaoyin. "Subjectively optimized HDTV video coding." (2009).

[5]. Shafique, Muhammad, Lars Bauer, and Jörg Henkel. "Optimizing the H. 264/AVC video encoder application structure for reconfigurable and application-specific platforms." Journal of Signal Processing Systems 60, no. 2 (2010): 183-210.

[6]. Kulikov, DrDmitriy, and Alexander Parshin. "Mpeg-4 avc/h. 264 video codecs comparison 2010-appendixes. MSU Graphics & Media Lab (Video Group), 2010."

[7]. Brandão, Tomás, Miguel Chin, and Maria Paula Queluz. "From PSNR to perceived quality in H. 264 encoded video sequences." proc. Of QoEMCS, Lisbon, Portugal (2011).

[8]. Prasantha, H. S., H. L. Shashidhara, K. N. B. Murthy, and M. Venkatesh. "Performance evaluation of H. 264 decoder on different processors." International Journal on Computer Science & Engineering 2, no. 5 (2010).

