Trends in Video Compression Technologies and Detailed Performance Comparison of H.264/MPEG-AVC and H.265/MPEG-HEVC

[1]Shivam Bindal, [2]Udit Khanna, [3]Manoj Sharma

[1, 2,3]Bharati Vidyapeeth's College of Engineering

New Delhi, India

Abstract- **In the reported work at hours has rigorously examined the video compression technologies and new trends for system on chip implementations.   The complexities involved in the integrating the complex video compressions algorithms for   an efficient solution targeting real time applications and storage applications pertaining to same are investigated. The advancements in recent years in the domain are reviewed highlighting the challenges and scope for further optimization sections involved.**

The objective of the paper is to provide an overview on past, present and future trends in Video Compression Technologies. We have reviewed the improvements and development in video encoding over the last two decade with future possibilities. Performance comparison of H.264/MPEG-AVC and H.265/MPEG-HEVC has also been surveyed and reviewed.

Index Terms-International Telecommunication Union-Telecommunications (ITU-T), Motion Picture Expert Group (MPEG), High Efficiency Video Coding (HEVC), H.264/MPEG-AVC, H.265/MPEG-HEVC.

I.   INTRODUCTION

The rapid development in Video Compression Encoding has been tremendous and there is substantial improvement in the video quality. Modern image and video compression techniques today offer the possibility to store and transmit vast amount of data necessary to represent digital images and videos in an efficient and robust way [1]. Digital image and video coding research started in the 1950s and 1960s with spatial DPCM (Differential Pulse Code Modulation)coding of images [2]. The first digital video coding standard "H.120" was developed in 1984 by ITU-T (International Telecommunication Union- Telecommunication) which had conditional replenishment and variable length coding. The second version was released in 1988 with improved motion compensation and background prediction. H.120 standard was not a widespread success and therefore is not in use these days.

Later in 1990, ITU-T developed a more enhanced video encoder and was standardized as H.261. Key aspects of H.261 were 16X16 macro block structure,

8X8 DCT (Discrete Cosine Transform), run-length, variable size length and scalar quantization.The second version of H.261 was released in 1993 which operated at 64-2048 kbps and featured backward compatible high resolution graphics trick mode.H.261 is still in use as backward compatibility feature in H.263.

Apart from ITU-T, there was ISO/IEC Moving Picture Expert Group (MPEG) which also dominated video encoding standardization. ISO/IEC developed MPEG-1 standards in 1993 which showed superior video quality when operated at higher bit rates due to bi-directional motion predictions i.e. B-frames, half pixel motion (1/2-pel), quantization weighting matrices etc. MPEG-1 provided approximately VHF quality using SIF 352X240/288 resolution [3].MPEG-1 helped in making possible digital audio broadcasting, digital cable/satellite TV and Video CDs.

In 1994, ISO/IEC Moving Picture Expert Group (MPEG) and ITU-T Video Expert Group (VCEG) formed a Joint Collaboration to develop H.262/MPEG-2 video compression standard. MPEG-2 was similar to MPEG-1 with additional support for interlaced videos and increased DC quantization precision. MPEG-2 was not optimized for bit rates less than 1Mbit/s. It outstripped MPEG-1 for bit rates greater than 2-3 Mbit/s. However, MPEG-1 forward compatibility is required for MPEG-2.

Video Compression technology took one step further in 1995 with the advent of H.263 by ITU-T VCEG with superior video quality at all bit rates. H.263 found its main application in Video Conferencing such as Flash video content due to its low bitrate compressed format. It was also used in 3GPP specification for IMS (IP Multimedia System) and Multimedia Messaging Service (MMS). The Baseline Algorithm [4] feature of H.263 superseded the H.261 with ½-pel motion compensation and 3D variable length coding of DCT coefficients. H.263 was further

developed as H.263+ or H.263 1998 (with improved compression efficiency of 15-20% in H.263+ over H.261) and H.263++ or H.263 2000. Fig 1 shows the timeline of the development of Video Compression Technologies from 1984 till present.
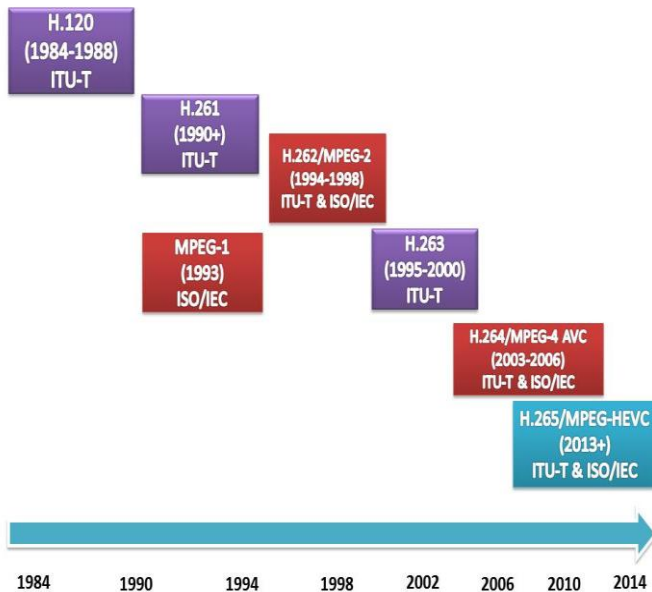

Fig 1.Chronology of Video Encoding Standards

This paper is organized as follows. In Section II and Section III, an overview of H.264/MPEG-AVC and H.265/MPEG-HEVC with their algorithms and System on Chip implementations has been given respectively. In Section IV, we describe a performance comparison of H.264/MPEG-AVC and H.265/MPEG-HEVCw.r.t System on Chip and the paper is concluded in Section V.

II. OVERVIEW OF H.264/MPEG AVC ALGORITHM AND SOC IMPLEMENTATION

The development of H.263++ and MPEG-4 'Visual' incorporated significant advancements in the domain of video coding, including error resilience, improved compression efficiency, zero-tree wavelet coding of still textures and coding of synthetic content [3]. However, their inability to deliver high quality video at relatively low bit rates and lack of network friendliness underscored the importance of a long term standardization activity. H.246 video coding spawned from a joint endeavor of the two leading organizations in video coding standards, namely, ITU-T and ISO/IEC. It was created, and is maintained, by the Joint Video Team (JVT) as a part of the Advanced Video Coding (AVC) project. Highlighting the capabilities and shortfalls of the H.264 standard and its subsequent advancements is a major part of this survey. We also specify a few areas of application that adapt this standard extensively.

The H.264/AVC encoder is collectively an integration of three basic processes that execute complementarily in the decoder to together result in achieving the key features of this standard, particularly,

- Effective motion compensation
- Reduced bit-rate for fixed fidelity
- Improved coding efficiency
- Good compression scheme.

The encoding process has been illustrated in Figure 2.

A. Frame Prediction

The prediction process is a combination of motion estimation and motion compensation. Every video frame is split into the basic processing unit called 'Macroblock', each of which is a 16x16 pixel matrix. For motion estimation, the encoder predicts Macroblocks on the basis of data that has been coded previously either in the current frame (spatial prediction), or other already transmitted frames (temporal prediction). A residual is then formed by subtracting this prediction from the current frame, in a process known as motion compensation [13].
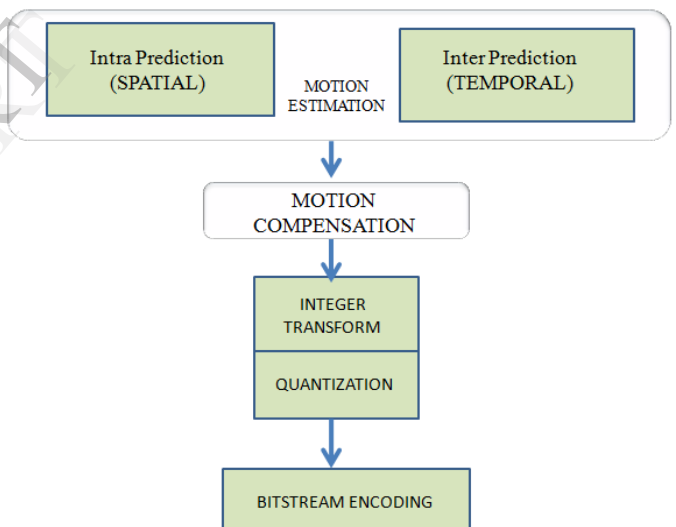

Fig 2. H.264/MPEG AVC Encoder Block Diagram

B. Transform and Quantization

Based upon a standard encoder-decoder basis pattern, weighting coefficients are obtained from residual samples. Integer transform is preferred to avoid erroneous inverse transformation in the decoder. To facilitate simple arithmetic, 16x16 Macroblock is divided into smaller 4x4 or 8x8 blocks. Between transformation and inverse transformation, quantization and re-scaling are introduced, to allow compression. Higher the quantization parameter, greater is the compression but the quality of decoded image decreases after re-scaling due to loss of some weighting coefficients in the matrix [13].

C. Bitsream Encoding or Entropy Coding

Certain values and parameters, known as syntax elements, regarding the structure of compressed data and quantized transform coefficients are necessary to be encoded along with frame-data to enable successful decoding of image. These values and parameters are stored or transmitted, after conversion into binary codes using VLC (variable length coding) and/or BAC (binary arithmetic coding). The standard under survey uses these modes in a Context Adaptive manner (CAVLC and CABAC).

D. Applications

The H.264 standard is adopted for a plethora of applications, out of which a few have been enlisted.

- High Definition DVDs (Blu-Ray formats)
- High Definition TV broadcasting
- Apple products including iTunes video downloads, iPod video and Mac OS
- NATO and US DoD video applications
- Mobile TV broadcasting
- Internet video
- Videoconferencing

Figure 3 shows a video frame compressed at the same bitrate using the H.264 standard and two of its predecessors.

Fig 3. A video frame compressed at the same bitrate using MPEG-2 (left), MPEG-4 Visual (center) and H.264 compression (right) [13].

III. OVERVIEW OF H.265/MPEG HEVC ALGORITHM AND SOC IMPLEMENTATION

High Efficiency Video Coding (HEVC) [11] is a video compression standarddeveloped by a Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T and ISO/IEC 23008-2. HEVC is a successor to H.264/MPEG-4 AVC and an evolution of the existing video coding recommendations (ITU-T H.261, ITU-T H.262, ITU-T H.263 and ITU-T H.264). H.265 was developed in response to the growing need for higher compression of motion picture for various ICT solutions like internet streaming, communication, videoconferencing, digital storage media and television broadcasting [5].The technical content specifications were officially finalized on April 13, 2013. The H.265 standard was designed to be applicable for almost all existing H.264 applications while putting emphasize on high resolution video [14].

HEVC particularly focuses on two key issues (1) increased video resolution (2) increased use of parallel processing architecture [6]. HEVC's key benefit is the reduction of bandwidth bitrate up to 50% while maintaining the same video quality.

A. Coding Tree Unit

In HEVC, picture is partitioned in coding tree blocks (CTBs) and uses a Quadtree based partitioning as shown in Fig 4.The size of the CTBs can be chosen by the encoder according to its architectural characteristics and the needs of its application environment, which may impose limitations such as encoder/decoder delay constraints and memory requirements [7]. Frames are split into Coding Tree Unit (CTU) which contains a luma CTB and 2 Chroma CTBs. CTB can be of the size 16X16, 32X32 or 64X64. More the size more will be the compression rate. CTU is a basic unit of the H.265 standard to specify decoding process. A 64X64 CTB is shown in Fig 5.
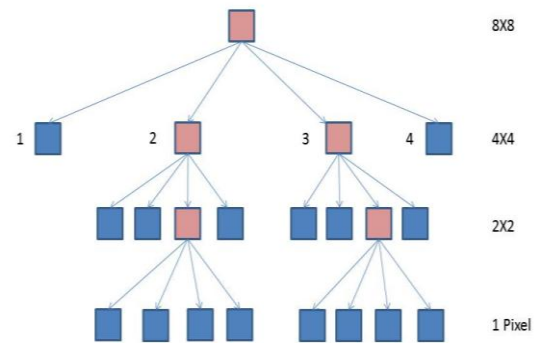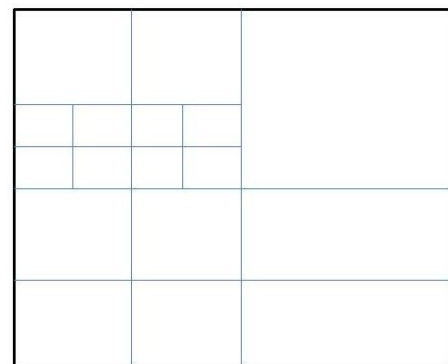

Fig 4.Quad Tree

Fig 5.CTB 64X64

The CTU may be split in four partitions in a recursive fashion down to coding as small as 8X8. This recursive split into four partitions is called a Quad tree [9].

B. Prediction Unit

There can be one, two or four prediction units in each Coding unit (CU). Therefore, total seven types of partitions are possible in HEVC. There are 35 intra-prediction modes including dc, planar and angular

modes in the Main profile. Prediction Unit may be predicted from either uni-prediction or bi-prediction reference location from up to 16 previously decoded frames [9].

C. Deblocking Filter

The HEVC draft standard defines two in-loop filters that can be applied sequentially to the reconstructed picture. The first one is the deblocking filter and the second one is the sample adaptive offset filter (SAO) that is currently included into the main profile [10]. Deblocking filter can use 4 input pixels on either side of the edge which lies on 8X8 grid resulting in filtering of adjacent edges. The filter is designed in such a way as to reduce the decoding complexity while improving the quality. In a typical architecture, the HEVC deblocking filter only consumes from 84 to 88 cycles per 16×16 block, which is less than half of the typical 200 cycles per 16×16 block cycle budget (for a 1080p@120 f/s video running at 250 MHz clock rate) [12].

A simplified HEVC video encoder is shown in Fig 6 with decoder modeling shaded in dark blue box.
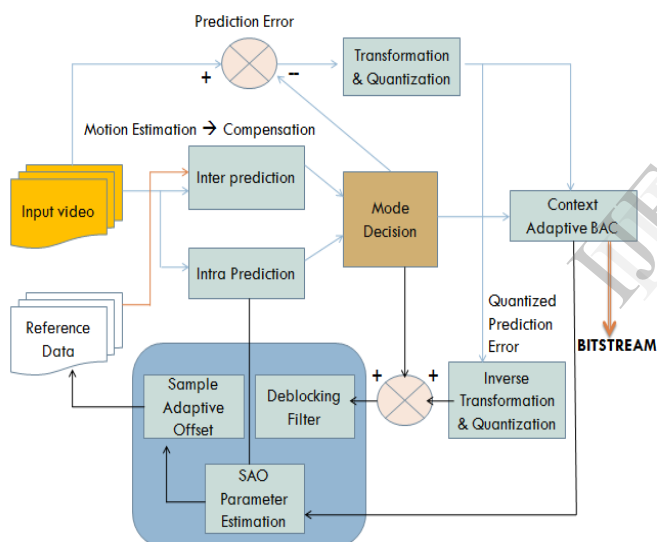

Fig 6.A Simplified HEVC Video Encoder.

D. Applications

HEVC termed as the future of encoding finds its application in numerous fields out of which few are listed below:

- 4KX2K and 8KX4K UltraHD (UHD) i.e. a resolution of up to 8192X4320.
- Video format bit rate conversion applications.
- High Definition (HD) TV Broadcasting.
- Videoconferencing.
- Internet Streaming.

IV. PERFORMANCE COMPARISON OF H.264/MPEG-4 AVC AND H.265/MPEG HEVC

A. Block Structure

There are substantial improvements in the block structure of H.265 over H.264. In H.264, every frame is divided in basic Macroblocks of 16X16 with intra blocks of 16 4X4 sub-blocks or one 16X16 block whereas in H.265 each frame is divided in CTB of 64X64, 32X32 or 16X16 (see Fig 5). Unlike H.264, H.265 has Quad tree sub-partitioning in coding blocks (see Fig 4). The quad tree partitioning allows for splitting into blocks of variable size according to the characteristics of the region covered by CTB. Also, in H.264 predictions and transforms are static whereas in H.265 it is flexible. Fig 7 shows the basic difference in the block structure of H.264 and H.265 standards.
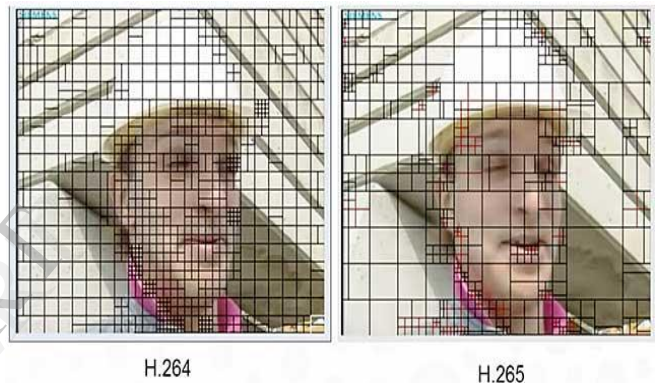

H.264   H.265

Fig 7.Difference in Block Structures of H.264 and H.265 [15].

B. Transformation Technique

Although the coding efficiency of the transforms depends on the statistical moments and probability distribution of the input signals [16], it is imperative to contrast the H.264 standard with its successor HEVC, on the basis of transformation approximation techniques used. While the H.264 uses a 4 x 4 Direct Cosine Transform over all block partitioning, HEVC allows flexible 'transform block' (TB) sizes of 4x4, 8x8, 16x16 and 32x32 [17].

C. Filters

Another qualitative comparison can be made on the basis of additional filtering introduced in HEVC. The In-Loop Deblocking Filter in HEVC assumes an adaptive index of 0 to 2, as compared to the scale of 0 to 5 in H.264. The additional sample-adaptive offset (SAO) aids to remove banding and sharpen edges. With each coding tree block, a look-up table is associated, whose parameters depend upon local gradient. The LUT attributes an offset addition to each sample.

D. Picture Classification Hierarchy

The HEVC standard has significantly modified and evolved the picture classification hierarchy. It provides an addition to it in the form of random access pictures (RAP), which include Clear Random Access (CRA) and Broken Link Access (BLA) pictures, besides the previously existing Instantaneous Decoder Refresh pictures. The pictures with a smaller display order and marked as 'Tagged for Discard' (TFD). For splice points in concatenated bit streams, the BLA type is used [17].
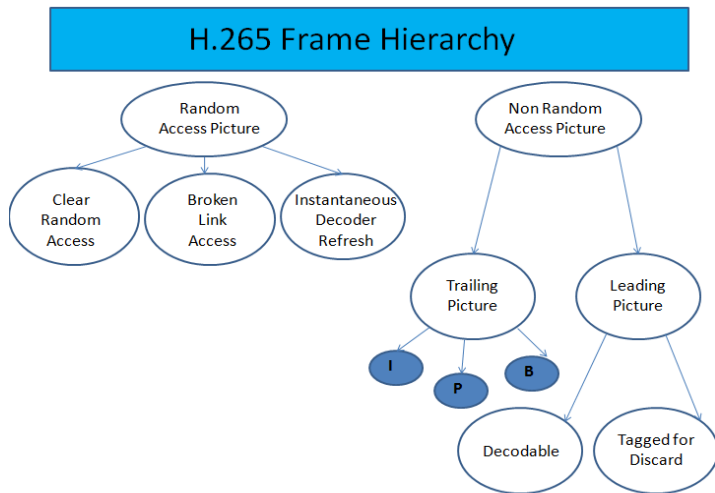

Fig 9. Picture Classification Hierarchy

E. Entropy Coding

Both the encoders are coded using CABAC i.e. Context Adaptive Binary Arithmetic Coding but coding of transform coefficient has been improved in H.265 using context selection schemes which is efficient for larger transform sizes. H.264 also uses CAVLC (Context Adaptive Variable Length Coding) which is faster but less efficient then CABAC.

F. Bit Rate

The major improvement in H.265 video encoding resulted in reduced bit rate of up to 50% maintaining the same/improved video quality as H.264 as can be seen from the Fig 8. The reduction in bandwidth up to half also reduces the cost of production. However, coding complexity increases in H.265 to achieve approx. 50% reduced bit rate.

V. CONCLUSION

Past, Present and Future of Video Compression was presented highlighting pros and cons of each w.r.t their architectures and present case scenario. In the last section a thorough performance comparison was made between H.264 and H.265 (which is summarized in Table 1) and comparison parameters clearly shows the improvements of H.265 over H.264 and how it can supersede H.264 in the coming decade.

Table 1.Summarized Comparison between H.264 and H.265.

Feature	H.264/AVC	H.265/HEVC	Remarks
1.Block Structure	Uses Macroblocks	Uses CTUs (Coding Tree Units)	H.265 uses quad tree partitioning of coding blocks which results in flexible coding tree structure.
2. CU Sizes	16X16 and 8X8	64X64, 32X32, 16X16 and 8X8	Improved coding efficiency in H.265. More complex than H.264.
3. PU Sizes	16X16, 16X8, 8X16, 8X8, 8X4, 4X8, 4X4	64X64, 64X32, 32X64, 32X32, 32X16, 16X32, 16X16, 16X8, 8X16, 8X8, 8X4, 4X8	More variation in Prediction Unit ensured greater coding efficiency with increased complexity in H.265.
4.TU Sizes	8X8, 4X4	32X32, 16X16, 8X8, 4X4	TUs contain coefficients for spatial block transform and quantization.
5. Intra Prediction Modes	9	35	H.265 uses Enhanced Hybrid Spatial-temporal prediction model
6. Deblocking Filter	✓	✓	Although present in both encoders, design is simplified for H.265 w.r.t decision- making and filtering process for parallel processing.
7. Sample Adaptive Offset (SAO)	✗	✓	SAO's goal in H.265 is to better reconstruct the original signal by using look up table.
8. Bit Rate Reduction	50% reduction from H.262/MPEG-2	40-50% reduction from H.264/MPEG-AVC	Bit Rate reduction is the major improvement of H.265 over H.264. Video quality is maintained at same level while reducing the bit rate and file size of the video.
9. Entropy Coding	CABAC/CAVLC	CABAC	Although CABAC is similar in both, it has gone under several improvements in H.265 like increased throughput speed and reduce context memory requirement.
10. Motion Vector (MV)	MVP is used.	AMVP is used.	Merge mode for MV can also be used in H.265.
11. Motion Compensation	6-tap filtering of half-sample position followed by linear interpolation for quarter sample position.	7-tap or 8-tap filters are used for interpolation of fractional sample position [8].	Better motion compensation in H.265 than H.264.
12. Ultra High Definition (UHD)	✗	Supports up to 8K (8192X4320) and up to 300fps.	H.265 promises to deliver the future of video encoding by supporting UHD resolutions.

REFERENCES

1. T. Sikora, *"MPEG digital video coding standard"*, IEEE Signal Process. Mag., vol. 14, no. 5, pp. 82–100, Sep. 1997
2. T. Sikora, "*Trends and Perspectives in Image and Video, Coding*", Invited Paper, IEEE.
3. Garry Sullivan, ITU-T, http://www.itu.int/ITU-T/worksem/vica/docs/presentations/S0_P2_Sullivan.pdf
4. http://marathon.csee.usf.edu/GaitBaseline/Algorithm.htm
5. "*Recommendation ITU-T H.265 High efficiency video coding*", http://www.itu.int/rec/T-REC-H.265
6. Davis P, SangeethaMarikkannan, "*Implementation of Motion Estimation Algorithm for H.265/HEVC*", International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering An ISO 3297: 2007 Certified Organization Vol. 3, Special Issue 3, April 2014, ISSN (Print) : 2320 – 3765 ISSN (Online): 2278 – 8875
7. Jens-Rainer Ohm, Gary J. Sullivan, Heiko Schwarz, ThiowKeng Tan, Thomas Wiegand, "*Comparison of the Coding Efficiency of Video Coding Standards—Including High Efficiency Video Coding (HEVC)*", IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, VOL. 22, NO. 12, DECEMBER 2012
8. Gary J. Sullivan, Jens-Rainer Ohm, Woo-Jin Han, Thomas Wiegand, "*Overview of the High Efficiency Video Coding (HEVC) Standard*", IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, VOL. 22, NO. 12, DECEMBER 2012 1649.
9. MehulTikekar, Chao-Tsung Huang, ChiraagJuvekar, Vivienne Sze, Anantha P. Chandrakasan, "*A 249-Mpixel/s HEVC Video-Decoder Chip for 4K Ultra-HD Applications*", IEEE JOURNAL OF SOLID-STATE CIRCUITS, VOL. 49, NO. 1, JANUARY 2014 61
10. AndreyNorkin, GisleBjøntegaard, ArildFuldseth, Matthias Narroschke, Masaru Ikeda, Kenneth Andersson, Minhua Zhou, Geert Van der Auwera, "*HEVC Deblocking Filter*", 1746 IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, VOL. 22, NO. 12, DECEMBER 2012
11. B. Bross, W.-J. Han, G. J. Sullivan, J.-R. Ohm, and T. Wiegand, *High, Efficiency Video Coding (HEVC) Text Specification Draft 8, ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11 document JCTVTC-J1003*, Joint Collaborative Team on Video Coding (JCTVC), Stockholm, Sweden, Jul. 2012.
12. M. Zhou, O. Sezer, and V. Sze, CE12 Subset 2: *Test Results and Architectural Study on De-Blocking Filter Without Parallel on/off Filter Decision, ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11*

document JCTVC-G088, Joint Collaborative Team on Video Coding (JCTVC), Geneva, Switzerland, Nov. 2011.

13. Prof. Iain Richardson, "*An overview of H.264 Advance Video Coding*", Available: http://www.vcodex.com/images/uploaded/469323879727520.pdf

14. Dan Grois, DetlevMarpe, AmitMulayoff, BenayaItzhaky, and OferHadar "*Performance Comparison of H.265/MPEG-HEVC, VP9, and H.264/MPEG-AVC Encoders*", PICTURE CODING SYMPOSIUM 2013 (PCS 2013), San José, CA, USA, Dec 8-11, 2013.

15. http://www.streamingmedia.com/Articles/Editorial/What-Is-.../What-Is-HEVC-(H.265)-87765.aspx

16. Narroschke.M," *Coding Efficiency of the DCT and DST in Hybrid Video Coding",* Selected Topics in Signal Processing, IEEE Journal of (Volume: 7, Issue: 6)**, Page** 1062 – 1071.

17. Andreas Unterweger, *"What is new in HEVC/"H.265"?"* October 17, 2012, http://dustsigns.de/CMS/wp-content/uploads/HEVC.pdf