Available online at www.elixirpublishers.com (Elixir International Journal)

Computer Science and Engineering

Elixir Comp. Sci. & Engg. 89 (2015) 37065-37069

ISSN: 2229-712X

An overview of H.26x series and its applications

S.Vetrivel
Mother Terasa College of Engineering, Thoothukudi , Tuticorin.

ARTICLE INFO

Article history:
Received: 11 December 2012;
Received in revised form:
15 December 2015;
Accepted: 21 December 2015;

Keywords
ITU, H.261,
H.262, H.263,
H.264, DCT,
CIF, QCIF,
AVC.

ABSTRACT

The advent of compression standards has led to the proliferation of innovative techniques in the compressed domain has become an active research topic. This paper presents an overview of the latest video compression standards related to the H.26xSeries and this paper is specifically covered including its latest standard, H.265 otherwise known as AVC (Advanced Video Coding). It also provides an overview of compression rules of thumb for different standards and the corresponding performance requirements for real-time implementations.

Introduction

Video compression is an essential enabler for Internet video streaming, digital video camcorders, cellular media, and personal video recorders applications and an increasing number of video codec (compression/decompression) industry standards and proprietary algorithms are available to make it practical to store and transmit video in digital form. The goal for image and video compression is to represent (or encode) a digital image or sequence of images in the case of video using as few bits as possible while maintaining its visual appearance [3].

In this paper, various video compression techniques are reviewed, starting from H.261 series. This paper starts with an explanation of the basic concepts of video codec design and then explains how these various features have been integrated into international standards, up to and including the most recent such standard, known as H.265/HEVC. Fig. 1 represents the progression of the ITU-T Recommendations and MPEG standards [3].

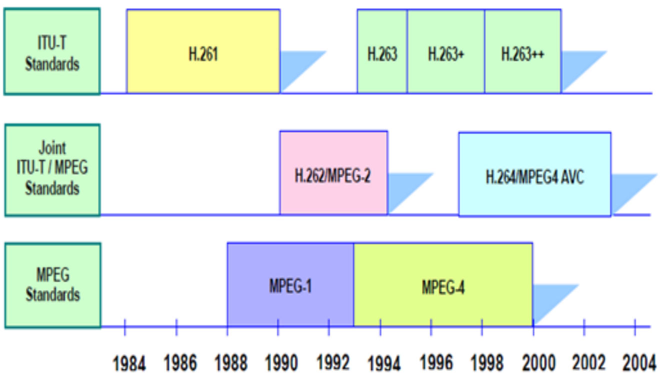

Fig. 1. Progression of the ITU-T Recommendations and MPEG standards

II. H.261 (1990)

H.261 is a codec designed by ITU (International Telecom Union) for video conferencing over PSTN (Public Switched Telephone Network). H.261 describes the video coding and decoding methods for the moving picture component of audiovisual services at the rate of p x 64 kbit/s, where p is in the range 1 to 30. It describes the video source coder, the video multiplex coder and the transmission coder.

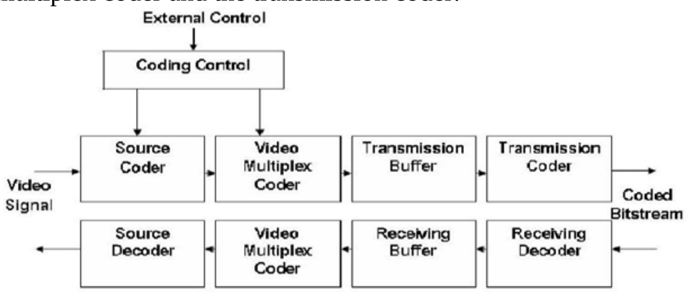

Fig. 2. H.261 Block diagram from ITU recommendation

Fig. 2 represents an overview of the H.261 CODEC, taken from the ITU reference documentation; shows the major components used to code and decode the bit streams [2]. H.261 encoding is based on the discrete cosine transform (DCT) and allows for fully-encoding only certain frames (INTRA-frame) while encoding the differences between other frames (INTER-frame). The main elements of the H.261 source coder are prediction, block transformation (spatial to frequency domain translation), quantization, and entropy coding. While the decoder requires prediction, motion compensation is an option. Another option inside the recommendation is loop filtering. The loop filer is applied to the prediction data to reduce large errors when using interframe coding. Loop filtering provides a noticeable improvement in video quality but demands extra processing power.

Fig. 3 shows an overview of the H.261 source coder, taken from the ITU reference documentation, shows the relationship between the DCT, prediction, and motion estimation logic flow [2].H.261 defines two picture formats: CIF (Common Intermediate Format) has 288 lines by 360 pixels/line of luminance information and 144 x 180 of chrominance information; and QCIF (Quarter Common Intermediate Format) which is 144 lines by 180 pixels/line of luminance and 72 x 90 of chrominance. Data for each macroblock consists of a

Tele:
E-mail addresses: vetri76@gmail.com

macroblock header followed by data for blocks [3]. Four luminance blocks and the spatially corresponding color difference blocks make up a macroblock which is shown in Fig. 4. Pictures are encoded as *luminance* (Y) and *two color difference components* (CB and CR). Y, CB and CR components are each function of the standard chrominance components (red, green, blue) and are defined in CCIR Recommendation 601 [8].

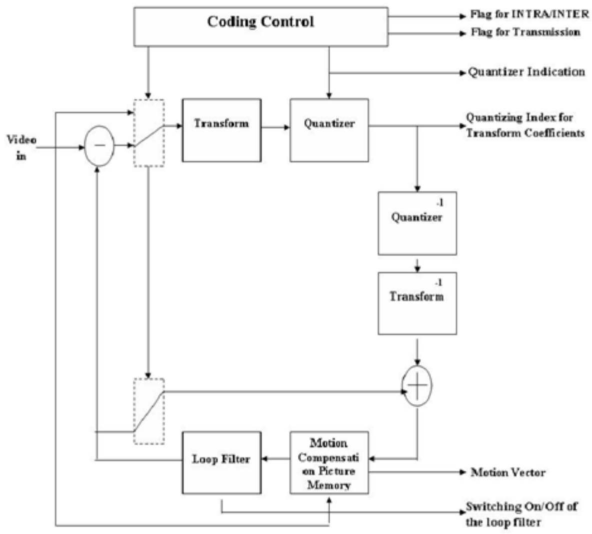

Fig. 3. H.261 Source coder block diagram

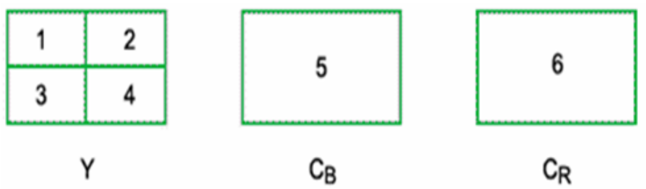

Fig. 4. Arrangement of blocks within a macroblock under H.261

A. Applications:

H.261 is the most important application in the video conferencing and video communications systems. Its uses include studio based video conferencing, desktop video conferencing, surveillance, monitoring, computer training, and tele-medicine. But it does not work well over Frame Relay or TCP/IP Internet as it is optimized only for low data rates and relatively low motion video. So, it led to the development of H.262.

III. H.263 (1995)

H.263 is a member of the H.26x family of video coding standards designed by ITU (International Telecom Union). H.263 was developed after H.261 with a focus on enabling better quality at even lower bitrates. One of the major original targets was video over ordinary telephone modems that ran at 28.8 Kbps at the time. The target resolution was from SQCIF (128x96) to CIF. The basic algorithms are similar to H.261 but with some features.

H.263 has been optimized for a large range of bit rates and not just 64K bits/s like with H.261. Another improvement is that it supports five resolutions. Besides the CIF (Common Interchange Format) and QCIF (Quarter Common Interchange Format) that H.261 supports, it also supports SQCIF, 4CIF, and 16CIF. 4CIF is 4 times the resolution of CIF, and 16CIF is 16 times the resolution. SQCIF is about half the resolution of QCIF. This means that H.261 compares well with the MPEG standards. Other features that make it similar to MPEG are Syntax-based Arithmetic Coding, Unrestricted Motion Vectors, and Advance Prediction Mode, along with Frame Prediction.

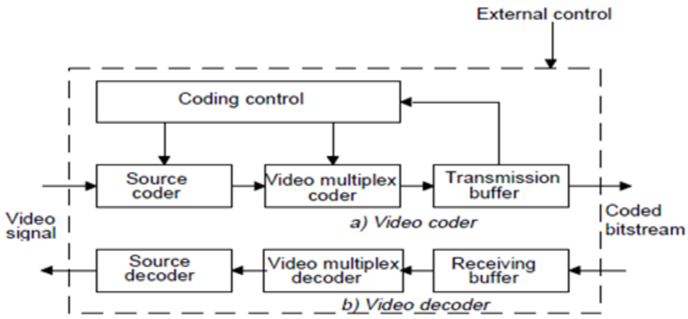

Fig. 5. Block Diagram of the H.263 Video Codec

The original H.263 had 8 extensions (or annexes): Inverse Transform Accuracy Specification, Hypothetical Reference Decoder, and Syntax-based Arithmetic Coding. *Versions:*

- Version 2 in 1997 known as H.263v2 or H.263+
- Improved Version in 2000 known as H.263v3 or H.263++
- In 2001 and 2005 some more modifications were made and the final version was published.

A. Applications:

An H.263 video can be played back on LGPL-licensed libav encoded VLC Media and MPlayer multimedia players, Real Video, Windows Media Player, and QuickTime.

IV. H.262/MPEG-2 Part 2 (1995)

H.262 or MPEG-2 Part 2 (formally known as ISO/IEC 13818-2), also known as MPEG-2 Video, is a digital video compression and encoding standard developed and maintained jointly by ITU-T Video Coding Experts Group (VCEG) and ISO/IEC Moving Picture Experts Group (MPEG). It is the second part of the ISO/IEC MPEG-2 standard. The ITU-T Recommendation H.262 and ISO/IEC 13818-2 documents are identical. Encoding for compressed video and audio data multiplexed with signaling information in a serial bistream.

A. Applications:

The format was initially developed to serve the transmission of compressed television programs via broadcast, cablecast, and satellite, and subsequently adopted for DVD production and for some online delivery systems. MPEG-4 lacks in in-loop deblocking filter, GMC being too computationally intensive, and OBMC being defined but not allowed in any profiles among other things.

V. H.264/AVC (2003)

H.264/AVC was finalized in March 2003 and approved by the ITU-T in May 2003 [6].An ITU standard for compressing video based on MPEG-4 that is popular, especially for high-definition video such as Blu-ray. Taking advantage of today's high-speed chips, H.264 delivers MPEG-4 quality with a frame size up to four times greater. It can also provide MPEG-2 quality at a reduced data rate, requiring as little as one third the original bandwidths.

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 [5]. It has three profiles. They are the *Baseline*, *Main*, and *Extended* profiles. It significantly outperforms the other codecs [7].

A. Video Coding Layer [9]

As in all prior ITU-T and ISO/IEC JTC1 video standards since H.261, the VCL design follows the so-called block-based hybrid video coding, in which each coded picture is represented in blockshaped units of associated luma and chroma samples called *macroblocks*. The basic source-coding algorithm is a hybrid of inter-picture prediction to exploit temporal statistical dependencies and transform coding of the prediction residual to exploit spatial statistical dependencies. There is no single coding element in the VCL that provides the majority of the significant improvement in compression efficiency in relation to prior video coding standards. It is rather a plurality of smaller improvements that add up to the significant gain.

Formerly known as "H.26L" by the ITU, H.264 is also known as "MPEG-4 Part 10" by the ISO MPEG group, which jointly developed the codec. Like most video coding standards, H.264 actually standardizes only the "central decoder...such that every decoder conforming to the standard will produce similar output when given an encoded bitstream that conforms to the constraints of the standard," according to Overview of the H.264/AVC Video Coding Standard published in *IEEE Transactions on Circuits and Systems for Video Technology* (ITCSVT). Basically, this means that there's no standardized H.264 encoder.

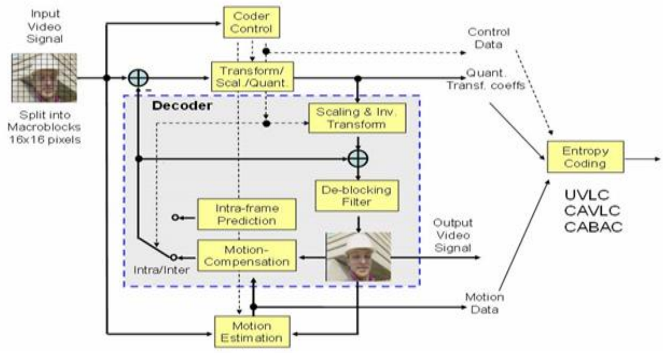

Fig. 6. Block diagram of typical encoding process for the VCL of H.264/AVC

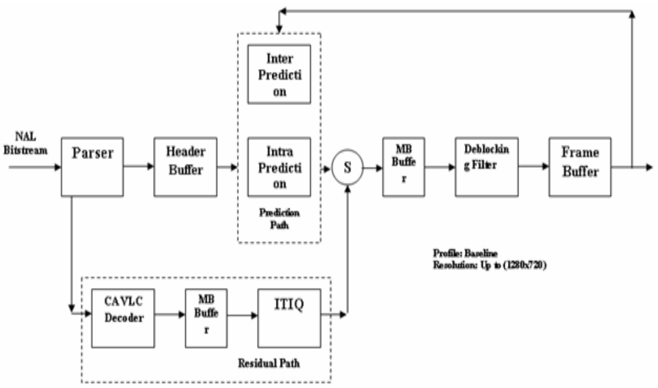

Fig. 7. H.264 Decoder block diagram

• ***Pictures, frames, and fields:*** A coded video sequence in H.264/AVC consists of a sequence of *coded pictures*. A coded picture in can represent either an entire *frame* or a single *field*, as was also the case for MPEG-2 video. Generally, a frame of video can be considered to contain two interleaved fields, a top and a bottom field. The top field contains even-numbered rows 0, 2,..., H/2-1 with H being the number of rows of the frame. The bottom field contains the odd-numbered rows (starting with the second line of the frame). If the two fields of a frame were captured at different time instants, the frame is referred to as an interlaced frame, and otherwise it is referred to as a progressive frame [5]. The more redundancy, the higher the quality at any given bit rate. To leverage this redundancy, H.264 streams include three types of frames

1.) **I-frames:** Also known as key frames, I-frames are completely self-referential and don't use information from any other frames. These are the largest frames of the three, and the highest-quality, but the least efficient from a compression perspective.

2.) **P-frames:** P-frames are "predicted" frames. When producing a P-frame, the encoder can look backwards to previous I or P-frames for redundant picture information. P-frames are more efficient than I-frames, but less efficient than B-frames.

3.) **B-frames:** B-frames are bi-directional predicted frames. As you can see in Fig. 8, this means that when producing B-frames, the encoder can look both forwards and backwards for redundant picture information. This makes B-frames the most efficient frame of the three. Note that B-frames are not available when producing using H.264's Baseline Profile.

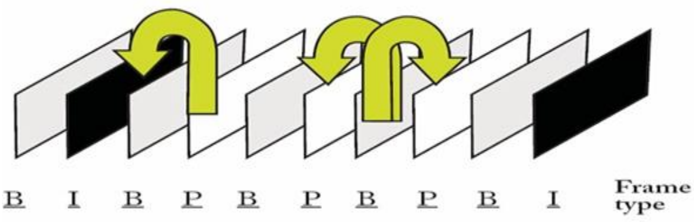

Fig. 8. Various frame types

• ***YCbCr color space and 4:2:0 sampling:*** The human visual system seems to perceive scene content in terms of brightness and color information separately, and with greater sensitivity to the details of brightness than color. Video transmission systems can be designed to take advantage of this. (This is true of conventional analog TV systems as well as digital ones.) In H.264/AVC as in prior standards, this is done by using an YCbCr color space together with reducing the sampling resolution of the Cb and Cr chroma information.

• ***Division of the picture into macroblocks:*** A picture is partitioned into fixed-size macroblocks that each covers a rectangular picture area of 16x16 samples of the luma component and 8x8 samples of each of the two chroma components. Macroblocks are the basic building blocks of the standard for which the decoding process is specified. The basic coding algorithm for a macroblock is described after we explain how macroblocks are grouped into slices.

• ***Slices and slice groups:*** Slices are a sequence of macroblocks which are processed in the order of a raster scan when not using flexible macroblock ordering (FMO) which is described in the next paragraph. A picture maybe split into one or several slices as shown in Fig. 9**.** A picture is therefore a collection of one or more slices in H.264/AVC. Slices are self-contained in the sense that given the active sequence and picture parameter sets, their syntax elements can be parsed from the bitstream and the values of the samples in the area of the picture that the slice represents can be correctly decoded without use of data from other slices provided that utilized reference pictures are identical at encoder and decoder [9]. In H.264/AVC, the macroblocks are processed in so called slices whereas a slice is usually a group of macroblocks processed in raster scan order [6].

• ***Encoding and decoding process for macroblocks:*** All luma and chroma samples of a macroblock are either spatially or temporally predicted, and the resulting prediction residual is encoded using transform coding. For transform coding purposes, each color component of the prediction residual signal is subdivided into smaller 4x4 blocks. Each block is transformed using an integer transform, and the transform coefficients are quantized and encoded using entropy coding methods.

• ***Adaptive frame/field coding operation:*** If a frame consists of mixed regions where some regions are moving and others are not, it is typically more efficient to code the non-moving regions in frame mode and the moving regions in the field mode. Therefore, the frame/field encoding decision can also be made independently for each vertical pair of macroblocks (a 16x32 luma region) in a frame. This coding option is referred to as macroblock-adaptive frame/field (MBAFF) coding [9]. A mixture of field and frame macroblock pairs may occur within an MBAFF frame, the methods that are used for

1) Zig-zag scanning
2) Prediction of motion vectors
3) Prediction of intra prediction modes
4) Intra frame sample prediction
5) Deblocking filtering
6) Context modeling in entropy coding is modified to account for this mixture.

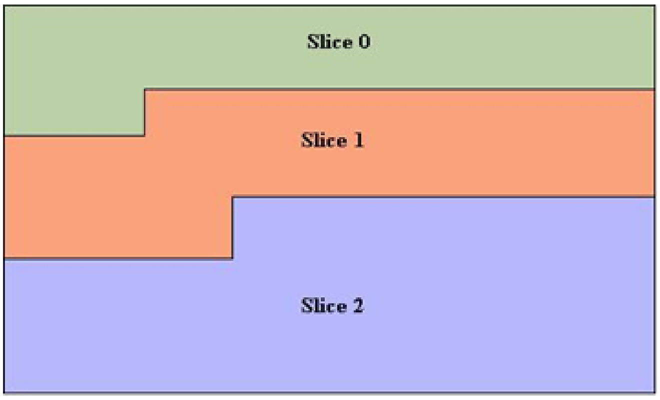

Fig. 9. Partitioning of an image into several slices

• ***Intra-frame Prediction:*** Each macroblock can be transmitted in one of several coding types depending on the slice-coding type. In all slice-coding types, the following types of intra coding are supported, which are denoted as Intra_4x4 or Intra_16x16 together with chroma prediction and I_PCM prediction modes.

• ***Inter-frame Prediction****:* Various Inter-frame Predictions are

1.) Inter-frame Prediction in P Slices
2.) Inter-frame Prediction in B Slices

• ***Transform, Scaling, and Quantization:*** Similar to previous video coding standards, 264/AVC
utilizes transform coding of the prediction residual [9]. However, in H.264/AVC, the transformation is applied to 4x4 blocks, and instead of a 4x4 discrete cosine transform (DCT), a separable integer transform with similar properties as a 4x4 DCT is used.

• ***Entropy Coding:*** In H.264/MPEG4-AVC, many syntax elements are coded using the same highly-structured infinite-extent variable-length code (VLC), called a zero-order exponential-Golomb code. A few syntax elements are also coded using simple fixedlength code representations [9]. Most of the video coding standards today, including H.264 and MPEG-2, are based on hybrid video coding-video is compressed using a hybrid of motion compensation and transform, coding. These video coding algorithms compress the video data by reducing the redundancies inherent in video, which fall into four classes: [4]

1.) Spatial
2.) Temporal
3.) Perceptual and
4.) Statistical.

• ***In-Loop Deblocking Filter:*** H.264/AVC defines an adaptive in-loop-deblocking filter, where the strength of filtering is controlled by the values of several syntax elements [9].

• ***Hypothetical Reference Decoder:*** One of the key benefits provided by a standard is the assurance that all the decoders compliant with the standard will be able to decode a compliant compressed video. To achieve that it is not sufficient to just provide a description of the coding algorithm. It is also important in a real time system to specify how bits are fed to a decoder and how the decoded pictures are removed from a decoder. Specifying input and output buffer models and developing an implementation independent model of a receiver achieves this. That receiver model is also called Hypothetical Reference Decoder (HRD) and is described in detail in. An encoder is not allowed to create a bitstream that cannot be decoded by the HRD. Hence, if in any receiver implementation the designer mimics the behavior of HRD, it is guaranteed to be able to decode all the compliant bitstreams.

In H.264/AVC HRD specifies operation of two buffers: [9]

1.) Coded Picture Buffer (CPB)
2.) Decoded Picture Buffer (DPB)

B. Applications and current status of Deployments [5]

As a generic, all-purpose video coding standard that is able to cover a broad spectrum of requirements from mobile phone to digital cinema applications within a single specification, H.264/MPEG4-AVC has received a great deal of recent attention from industry. Besides the classical application areas of videoconferencing and broadcasting of TV content (satellite, cable, and terrestrial), the improved compression capability of H.264/MPEG4-AVC enables new services and thus opens new markets and opportunities for the industry.

As an illustration of this development, consider the case of "mobile TV" for the reception of audio-visual content on cell phones or portable devices, presently on the verge of commercial deployment. Several such systems for mobile broadcasting are currently under consideration,
e.g.,

• Digital Multimedia Broadcasting (DMB) in South Korea
• Digital Video Broadcasting — Handheld (DVB-H), mainly in Europe and the United States
• Multimedia Broadcast/Multicast Service (MBMS), as specified in Release 6 of 3GPP

It needs a higher hardware costs tend to out-weigh the benefits in a practical environment and it requires higher speed grade processors for encoding powerful PC to decode. It is limited yet growing to support VMS. So, it led to the next series.

V. H.265/HEVC

The Moving Picture Experts Group (MPEG) and Video Coding Experts Group (VCEG) are examining the need for a new video compression standard. At a recent MPEG meeting (June 2009), several proposals for improved video compression were presented. The consensus was that (a) there is likely to be a need for a new compression format, as consumers demand higher-quality video and as processing capacity improves and (b) there is potential to deliver better performance than the current state-of-the art.

Table 1. Picture Formats Supported by H.263

Luminance pixels	Luminance lines	H.261 support	H.263 support	Uncompressed bitrate (Mbit/s)			
				10 frames/s		30 frames/s	
				Grey	Color	Grey	Color
128	96		Yes	1.0	1.3	3.0	4.4
176	144	Optional	Yes	2.0	3.0	6.1	9.1
704	576		Optional	8.1	12.2	24.3	36.5
1408	1152		Optional	129.8	194.6	389.3	583.9

The current plan is to set up a Joint Collaborative Team (JCT) of MPEG and VCEG representatives to work on a new video coding standard. It will aim to deliver significantly better compression performance than H.264/AVC, probably at a higher computational cost and reduce bitrate requirements by half with comparable image quality.

A. Applications:

HEVC is targeted at next-generation HDTV displays and content capture systems which feature progressive scanned frame rates and display resolutions from QVGA (320x240) up to 1080p and Super Hi-Vision, as well as improved picture quality in terms of noise level, color gamut and dynamic range.

VI. Conclusions and Future work

This paper has presented an overview of the H.26x Series of various compression techniques. It has reviewed several existing video compression techniques such as H.261, which was designed by ITU for video conferencing over PSTN which supports low data rates and relatively low motion video, where H.263 came into existence which focus on enabling better quality at even lower bitrates. H.262/MPEG-2 Part 2 which is rarely in usage. Then the H.264 video coding standard is gaining momentum and represents a number of advances in terms of both coding efficiency enhancement and flexibility for effective use over a broadband variety of network types and application domains. There is no standardized H.264 encoder. And now the Emerging H.265/HEVC is used to reduce bitrate requirements by half with comparable image quality by approximately July 2012. The future work will be done on bit rate reduction to improve the QOS (quality of service) by reducing the processing power and most significantly the time delay of transport for sensitive applications.

References

[1] Al Tamimi A.K.; Jain R.: Modeling and Prediction of High Definition Video Traffic: A Real-World case Study and So-In, C.

[2] Citeseer, J.C; Oliveira de: A Java H.263 Decoder Implementation.

[3] Golston, J: Comparing media codecs for video content”, Embedded Systems Conference, San Francisco.

[4] Kalva, H. (2006): The H.264 video coding standard, IEEE multimedia, 13(4), pp. 86--90.

[5] Marpe, D. and Wiegand, T. and Sullivan, G.J. (2006): The H.264/MPEG4 advanced video coding standard and its applications: IEEE Communications Magazine, 44(8), pp. 134--143.

[6] Ostermann, J.; Bormans J.; List P.; Marpe D.; Narroschke M.; Pereira F.; Stockhammer T.; Wedi T.:Video coding with H.264/AVC.

[7] Sullivan, G.J. Wiegnad, T. (2005): Video compression—from concepts to the H.264/AVC standard. proceedings of the IEEE 93(1), pp. 18--31.

[8] Turletti, T. (1993): H.261 software codec for videoconferencing over the Internet.

[9] Wiegand, T; Sullivan, G.J; Bjontegaard, G; Luthra, A. (2003): Overview of the H.264/AVC video coding standard: IEEE Transactions on circuits and systems for video technology, 13(7), pp. 560--576.