International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2015): 78.96 | Impact Factor (2015): 6.391

The Evolution of Video Codecs and the Future

Sharath Guna

Student of Computer Science and Engineering Department, PES Institute of Technology, South Campus, Electronic City, Bangalore 560021, India

Abstract: ***The evolution of video codecs by various experts group and their contribution to the global media is highly noteworthy. This paper gives a study about the evolution of various video codecs and the comparison between them along with the insights to the future of video codecs.***

Keywords: codecs, video compression, x264, hevc

1. Introduction

One of the main problems in conveying the source data is maintaining an acceptable fidelity within a constrained bit-rate. Also it may be posed as conveying the source data using the lowest bit rate possible whilst maintaining the fidelity. The coding system which performs this operation is called a Codec. Since the advent of the first codec system, there have been drastic changes in the way a file is coded (-decoded) and the efficiency has increased over the timeline.

2. What is a CODEC?

A Codec in its entirety is a computer program or device which has the ability to encode or decode digital streams of data or signal.

CO-DEC = COder + DECoder or, less commonly known as "compressor-decompressor".

A codec encodes a digital data stream or signal for transmission, or decodes it for playback or editing. A codec should not be confused with a compression format which is a standard, while a codec is a tool which is capable of reading or writing such files. Codecs find applications in real-time video-conferencing, online streaming media and video-editing etc. [4]

Codecs such as MPEG and H.26x series paved the way for present developments of digital television and storage without which it would have been impossible taking into account, the bandwidth capacity. Certain standard authorities like ISO and ITU standardize the codecs which provide a stable environment for broadcasters and manufacturers to develop systems in accordance.

3. Background

Raw digital video before compression requires a very high transmission rate in the range of few hundred Mb/s or equivalently a storage capacity of many hundreds of Gbs per 2 hr movie. However practical channel bandwidth is respectively limited to about 13-30 Mb/s for terrestrial transmission, 40 Mbit/s for satellite and cable, and 4-9GB for storage devices. In order to achieve this reduction in bit rate between studio and broadcast quality video requires compression prior to transmission and recompression at the receiver. A standard Codec is employed to achieve this process. [2]

Table 1: Illustrative comparison ratios for production television applications at recommended qualities [2]

Pixel resolution	Format	Input rate from studio Mbits/s	Output rate in Mbit/s	Compression Ratio
1920 x 1080	HDTV	829	20	41:1
720 x 5576	SDTV	166	3-6	33:1
360 x 288	SIF	31	1	31:1

Fortunately, subsequent video frames are usually redundant whilst creating the perception of a video by moving images. It is this redundancy and others that are exploited to achieve the compression of the video.

Over the last two decades, there has been a tremendous change with continual evolution in video codecs and there is likely no sign that this evolution will come to an end. MPEG-2, a widely standardized codec during the 1990s was overtaken with respect to the degree of achievable compression by H263,MPEG-4 and in 2003 by H.264, also by the most recent HEVC. Along with H.261, MPEG-1 and MPEG-2 were the first codecs that combined multiple ways using various algorithms of removing redundancy in one codec. Basically, each video frame is split into blocks and matching blocks between successive frames are sought. Only the difference after it has been further encoded is then transmitted or stored.

One of the latest widely used codec to emerge, H.264, has taken advantage of the hardware, as the achievable computational complexity has increased in line with Moore's law .The size of the blocks that are compared has been reduced and made more flexible, thereby reducing the difference data that remains to be encoded. [2]

4. How Video Compression Works?

Video compression is all about the reduction and removal of redundant video data so that a digital video file can be effectively sent and stored. This process involves applying algorithms to the source video to create a compressed file that is ready for transmission or storage. To play the compressed file, an inverse algorithm is applied to produce a video that essentially shows the same content as the original source video. The time taken to compress, send, decompress and display a file is called latency. The more advanced the compression algorithm, the higher is its latency. A pair of algorithms that works together is called a video codec (encoder/decoder). Video codecs that make use of different standards are normally not compatible with each other; means that a video content that is compressed using one standard cannot be decompressed with a different standard. For instance, an MPEG-4 Part 1 decoder will not work with an H.264/AVC encoder. This is simply because one algorithm cannot correctly decode the output from another algorithm but it is possible to implement many different algorithms in the same software or hardware, which would then enable multiple formats to be compressed.

Different video compression standards utilize different methods of reducing data, and hence, results differ in bit rate, quality and latency etc. Results from encoders that use the same compression standard may also vary because the designer of an encoder can choose to implement different sets of tools defined by a standard. As long as the output of an encoder conforms to a standard's format and decoder, it is possible to make different implementations. This is advantageous because different implementations have different goals and budget. Professional non-real-time software encoders for mastering optical media should have the option of being able to deliver better encoded video than a real-time hardware encoder for video conferencing that is integrated in a hand-held device. A given standard, therefore, cannot guarantee a given bit rate or quality.

Furthermore, the performance of a standard cannot be properly compared with other standards, or even other implementations of the same standard, without first defining how it is implemented. A decoder, unlike an encoder, must implement all the required parts of a standard in order to decode a compliant bit stream. This is because a standard specifies exactly how a decompression algorithm should restore every bit of a compressed video. [4]

5. Evolution of Video Codecs

In order to understand the evolution of video codecs it is necessary for us to understand the standardization process. Due to engineering research there is a continual invention or refinement of compression algorithms, which is reported in journals such as those of the IEEE in the US. These innovations, after competitive assessment, are encapsulated by one of the two standards bodies, the ISO and the ITU, in standard codec specifications such as the MPEG and the H.26x series. However, the standard body standardizes only the format of the bit stream arriving at the decoder end, though obviously it is aware of algorithms that can exploit the information in the bit stream. The advantage of this procedure is that successive refinements can be made to the algorithms at the encoder or sender side, without occasioning the replacement of end-users' equipment such as the set-top boxes, digital televisions, and so on.

Therefore, successive refinements can take place in the lifecycle of a codec such as motion estimation, noise reduction, and advanced pre-processing.

In MPEG-2's case the rate of improvement has declined according to an exponential rule. That is the most gains were made in the first 3-4 years, whereas later improvements have not reduced the bit rate by as much. In fact, from about 2002 improvements have bottomed out or rather have approached an asymptote. Well before reaching the asymptote a new codec (MPEG-4) was introduce. At this point, there is said to be a Leap Gain in compression efficiency, when a new codec by its structure is able to take advantage not only of all the existing improvements from a previous codec but also those innovations that have been stored up, as it were, in the research literature. An alternative coding scheme known as H.263 predates MPEG-4 and has a similar performance. In H.264 the two standards tracks have merged. While the new and old codec continue to improve over time the new codec continues beyond the effective life-time of the old, which has reached its asymptote. [2]

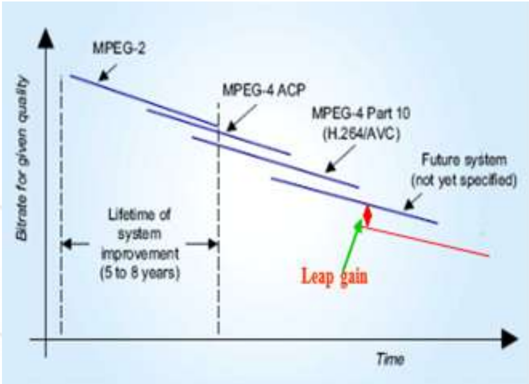

Figure 1: Stylized evolution of the video codecs over time.

New codecs are being developed to service new applications. MPEG-2 was developed for video broadcast, whereas its predecessor was intended for video storage on CD-ROM. H.263 was intended for video conferencing. From MPEG-4, the MPEG series have diverged towards compression services, including video animation and video database construction. H.264 aims to serve a variety of from very low bit rates of less than 20 kbit/s to HDTV quality video at around 20 Mbit/s. [2]

Our prediction on year-on-year reduction in bit rates is closer to around 7%, based on a more considered estimation of codec life-cycles, an expectation of higher video quality in the future, and higher performance from video decoders and displays. The principle reason for the retention of a codec is the need to sustain the economic life of consumer devices and video services. However, an early codec such as H.261 is still retained for some services for reasons of backwards compatibility and to maintain video conferencing across legacy ISDN circuits which is a precursor of 'broadband'. [2]

6. Various Codecs

JPEG For single-frame image compression, the industry standard with the greatest acceptance is JPEG (Joint Photographic Experts Group). JPEG consists of a minimum implementation (called a baseline system) which all implementations are required to support, and various extensions for specific applications. JPEG has received wide acceptance, largely driven by the proliferation of image manipulation software which often includes the JPEG compression algorithm in software form as part of a graphics illustration or video editing package. The image frame consists of three 2-D patterns of pixels, one for luminance and two for chrominance. Because the human eye is less sensitive to high-frequency color information, JPEG calls for the coding of chrominance (color) information at a reduced resolution compared to the luminance (brightness) information. In the pixel format, there is usually a large amount of low-spatial-frequency information and relatively small amounts of high-frequency information. [1]

H.120 was the first video-coding standard developed by CCITT (now ITU-T) in 1984 for video conferencing applications. This codec was based on Differential Pulse Code Modulation (DPCM), scalar quantization, and conditional replenishment (direct re-use of similar regions from a prior frame). It supported bit rates of 1.544 and 2.048 Mbit/s of the first digital hierarchy of North America and Europe respectively. This codec was abandoned soon afterwards and is not in use anymore, as shortly afterwards a new standard H.261 was developed. [2]

H.261, developed in 1990 to replace H.120, is regarded as the basis or originator of all modern video compression standards. The basic structure (the hybrid coding structure) proposed in H.261 is still dominant today. H.261 was based on 16×16 MB motion estimation/compensation, 8×8 DCT, zigzag scanning of DCT coefficients, scalar quantization of those coefficients, and subsequent variable length coding (VLC). The other key aspects of this coder were a loop filter (to remove artifacts at block boundaries), integer-pixel motion compensation accuracy (optional) and 2-D VLC for coding of coefficients. H.261 operates at bit rates of k×64 kbit/s, where k is an integer with values in the rage from 1 to 30 and 64 kbit/s is the base rate for ISDN links. H.261 is still in use (mostly as a backward compatibility feature) but it has now been overtaken by H.263. [2]

MPEG-1 (1991) was the first coding standard for motion pictures developed by ISO. It was mainly developed for video storage applications (on CD-ROM). MPEG-1 utilizes the same structure as H.261 but introduces the concept of bi-directional prediction (B-pictures are predicted from anchor I- and P-pictures). MPEG-1 provides superior quality to H.261 when operating at high bit rates (>= 1 Mbit/s for CIF 360 ×288 pixels – spatial resolution). MPEG1 also adds half-pixel motion estimation to the H.261 design. [2]

MPEG-2 (also known as ITU-T H.262) was developed jointly by ISO and the ITU-T in the period 1994/95. It is one of the most commonly used video coding standards deployed at this time, particularly for DVD and Digital Video Broadcasting (DVB). MPEG-2 supports two new features namely: interlaced scan pictures and scalability. Otherwise, in all other aspects it is essentially the same as MPEG-1. Although MPEG-2 has many applications, it was designed mainly for high quality video at bit rates in the range 2-20 Mbit/s, and is not suitable for low-bit rate applications (below 1 Mbit/s). Its various applications are defined under levels and profiles.

H.263 was first developed in 1995 to replace H.261 as the dominant video conferencing codec, owing to its superior performance at all bit rates. In particular, at very low bit-rates it reduces the bit-rate by a factor of two compared to H.261. The basic algorithm in H.263 employs: half-pixel motion compensation; 3-D VLC of DCT coefficients; and median motion vector prediction. In addition, H.263 proposes many optional enhanced modes such as: increased motion vector range; advance prediction mode (with Overlapped Block Motion Compensation (OBMC) to counter blocking, and switching between one and four motion vectors (MVs)); optional arithmetic entropic coding; and coding of PB frames – two P and B pictures are coded as one unit, reducing overhead at low bit-rates. H.263 went through many refinement phases resulting in H.263+ (1998) and H.263++ (2000). Most of the improvements involved error resilience and scalability aspects to cater for a new range of applications over mobile networks and the Internet. [1]

MPEG-4 (v.1, early 1999) approximately follows the H.263 design and includes all prior features, including various trick modes (simulation of VCR functions such as fast replay). However, MPEG-4 can also code multiple objects within a video frame, with shape coding of video objects being an important object coding technique. MPEG-4 also includes zero-tree wavelet coding of still pictures, as well as dynamic 2D mesh coding of synthetic objects and facial animation modeling. There are many application profiles and levels in MPEG-4. Some of them are implemented while others are still at a development stage or may never be implemented. Therefore, MPEG-4 is better regarded as a toolset of compression tools, rather than a codec in the mould of MPEG-1/2. It has a high degree of complexity compared to MPEG-2. Version 2 of MPEG-4 introduces quarter-pixel motion compensation and global motion compensation. Despite many fanfares, MPEG-4 has not been as popular with manufacturers as anticipated. The demise of MPEG-4 can be ascribed to the failure of hardware manufacturers to take up its object-based features, as these would require a radical new design compared to the macro-block-based processing streams that manufacturers have been accustomed to. [1]

H.264/ Advanced Video Codec (AVC) (also known as MPEG-4 Part-10 or Joint Video Team (JVT), after the developers) is a state-of-the-art video codec, standardized in 2003. It is suitable for a wide range of applications such as broadcast with set-top-boxes, DVD storage, packet networks, and multimedia telephony systems. H.264 encompasses the full range of bit rates and quality resolutions, unlike some previous codecs. Profiles such as the High profile for HDTV and Blue-ray disc storage support a wide set of applications. The Baseline profile is intended for applications with limited computing resources, such as video-conferencing and mobile applications. The Main profile was intended for broadcast TV and storage but has been overtaken by the High profile. The Extended profile is intended for streaming applications, with robust coding, trick modes, and server switching. H.264 is the first video codec that has been explicitly designed for fixed-point implementation and the first network friendly coding standard. It has higher computational complexity but better coding efficiency than previous standards. H.264's better coding efficiency is mainly attributed to some of the features arising from H.263++, such as predictive intra-frame coding, multi-frame and variable block size motion compensation, quarter to one-eighth pixel motion estimation precision, integer DCT transform, adaptive in-loop de-blocking filtering, and more efficient/advance entropy coding. It has an increased range of quantization parameters, and employs Lagrangian optimized rate control. In the latter, the components of the target bit-rate in a rate-distortion function are divided between the individual coding parts in such a manner that maximum reduction in distortion is achieved at the expense of minimal increase in the bit rate. [2]

7. The Future

The emerging of a more efficient future generation video coding standard is of a high demand at the moment. There seem to be two main contenders for the position of the next state-of-the-art video compression standard: JCT-VC H.265/HEVC and Google VP9. The main aim of HEVC is to achieve twice more efficient compression compared to its predecessor H.264/AVC, and VP9 was developed to get half the bit-rate of its predecessor VP8 with royalty-free video codec. Intra compression is one of the main features that determine the compression efficiency of the whole codec. HEVC is being developed by JCT-VC group - the creators of AVC. It is an evolution of AVC concepts with some added innovations. On the other hand, VP9 is a Google initiative to get a royalty-free compression standard with efficiency that is superior to AVC. It expands techniques used in AVC and VP8 and is very likely to replace AVC at least in the Online streaming video services. [5]

Both HEVC and VP9 video compression standards are hybrid block-based codecs that rely on spatial transformations. The input video frame is initially partitioned into blocks of the same size called macroblocks. The compression and decoding process works within each macroblock. A macroblock is further subpartitioned into smaller blocks to perform prediction. There are two basic types of prediction: intra and inter. Intra-prediction works within a current video frame and is based upon the compressed and decoded data available for the block being predicted. Inter-prediction is used for motion compensation: a similar region on the previously coded frames close to the current block is used for prediction. The aim of the prediction process is to reduce data redundancy and, therefore, not storing excessive information in coded bitstream. Once the prediction is done, it is subtracted from the original data to get residuals that should only be compressed. Residuals are subject to Forward Discrete-Fourier Transform (DFT). DFT translates spatial residual information into frequency domain. Thus the remaining spatial redundancy of this information is partly reduced. Quantization is applied to the transformed matrix to lose insufficient information. The insufficiency threshold is predetermined by encoder configuration. The remaining data and the steps applied are subject to entropy coding, which makes it possible to get compressed bitstream. For inter- as well as intra-prediction purposes the compressed data should be restored in the encoder. The only data loss takes place after integer DFT and quantization. Dequantization and inverse DFT are performed to restore residuals. Then the restored residuals and the predicted values are summed up to get restored pixel values, identical to those achieved in the decoder. These restored values are used for intra-prediction within the current video frame. An additional frame post-processing stage is optionally applied to eliminate image blockiness introduced by DFT and quantization. The final restored and post-processed video frame is stored in Decoded Picture Buffer (DPB) for inter prediction of further frames. VP9 and HEVC both utilize the described general compression dataflow, but differ in details. [6]

8. Conclusion

Recent studies have reported that in the comparison of codecs x264, VP9, and x265 using clips from around 500 movies using 6 different quality metrics and found that both VP9 and x265 have 40–50% better quality at 1080p than x264. It stated that with the VMAF metric, which closely resembles human viewing perception, x265 performed substantially (19% to 22%) better than VP9. [5]

Both VP9 and HEVC provide higher compression efficiency compared to the current video compression standard AVC. HEVC provides better compression rates than VP9; however VP9 is patent-free and can be used without licensing expenses.

One of the main reasons for HEVC to be more efficient in intra coding compared to VP9 is the more angular intra prediction mode which provides the most significant influence on intra compression efficiency (about 7.7% bit rate savings). Finally the modified HM encoder with 10 intra prediction modes and without SAO post processing is still about 7% more efficient compared to the VP9 encoder. [5]

Hence, apart from the fact that HEVC is royalty encumbered, it is surely the best available codec at present. Its hands down HEVC.

References

[1] Dr.Saroj Choudhary, Purneshwari Varshney, "A Study of Digital Video Compression Techniques," ISSN-2250-1991, PARIPEX - Indian Journal of Research. (journal style)
[2] Prof. M. Ghanbari, Prof. D. Crawford, Dr. M. Fleury, Dr. E. Khan, Dr. J. Woods, Mr. H. Lu, Mr. R. Razavi, "Future Performance Of Video Codecs,", SES2006-7-13.
[3] Maxim P. Sharabayko, Oleg G. Ponomarev, Roman I. Chernyak, "Intra Compression Efficiency in VP9 and HEVC,".
[4] Deepak Dembla, Biren Patel, Ashish Kumar, Dr. Yogesh Bhomia, "Comparison of H.264 and MPEG-4 Codec Based on PSNR Peak Signal to Noise Ratio Algorithm," ISSN-2277-128X, International Journal of Advanced Research in
[5] Computer Science and Software Engineering, Volume 3, Issue 3, March 2013.
[6] "High Efficiency Video Coding, ", wikipedia.com, [Online]. Available:

"https://en.wikipedia.org/wiki/High_Efficiency_Video_Coding". (General Internet site)
[7] "VP9," wikipedia.com,[Online]. Available: https://en.wikipedia.org/wiki/VP9,, (General Internet site).

Author Profile

Sharath Guna is currently pursuing the B.E. degree in Computer Science and Engineering from PES Institute of Technology, South Campus. During 2014-2017, he was widely interested in Image and Video Processing. He is highly absorbed in doing research about Computer Vision. He is also interested in working on an open source video codec development across multi-platforms.