International Journal of Emerging Trends & Technology in Computer Science (IJETTCS)
Web Site: www.ijettcs.org Email: editor@ijettcs.org, editorijettcs@gmail.com
Volume XX, Issue XX, May-June 2022 ISSN 2278-6856

Survey on Video coding standards

Mahesh Pawaskar[1], Gaurav Vijay[2], Atul Deshpande[3]

[1]School of Engineering, Career Point University, Kota, Rajasthan, India,

[2] School of Engineering, Career Point University, Kota, Rajasthan, India,

[3]A. P. Shah Institute of Technology, Thane,
University of Mumbai, Maharashtra, India

Abstract: ***Digital video storing, video streaming and video conferencing are globally important industry, which will continue to spread across the networks, businesses, and homes. Since 1970, number of video and image coding techniques have been explored. Two international standardization bodies namely, Video Coding Expert Group (VCEG) of ITU-T and MPEG of ISO/IEC have ratified series of video coding standard. This paper reviews various standards of video coding, their features and applications. The rate-distortion performance of modern video compression scheme is the result of intra-picture prediction techniques, interaction between motion representation techniques, interaction between motion representation technique etc. Designer of video service has to select appropriate video coding scheme, which meet requirement such as efficiency, complexity, rate-distortion, flexibility etc. This paper starts with explanation of basic concepts of video codec design and then explain number of video coding standard up to most recent standard Versatile Video Coding (VVC).***

Keywords: MPEG standard, H.26x standard, Compression ratio, bit rate, Joint Video Team (JVT), Video Coding Expert Group (VCEG)

1. INTRODUCTION

Traditional analogue video is replaced with digital video because of its several advantages. Digital videos are compatible with that of the text and voice data. There is a need of efficient video coding standard to make storage and transmission of video information effectively. Video coding is to be prior to transmission and storage. Digital videos consume large amount of data, and if they are not compressed properly, it would be highly difficult to store and transmit video data. Although todays, data storage capacity, network bandwidth and computer power have been increased tremendously, demands of better quality of video are never stopped. If video format of 352x288 pixel with 24-bit color depth is to be transmitted at frame rate 25 fps, it would take nearly 60 Mbps bandwidth to transmit over internet, whereas it is possible to store only 10 seconds of video in a Compact Disc (CD). There are number of applications where there is need of video compression. Some of the major applications are given below [1].

- Storage format such as Digital Versatile Disk (DVD), personal video recorder etc.
- Wireline and wireless real time conversational service.
- Local Area Network (LAN) or Internet video streaming using Internet Protocol (IP) or Real Time Protocol (RTP)
- Broadcast, cable, pay-per-view services over satellite, subscription and terrestrial channel.

Transmitting compressed bit stream over the network is decisive. In case, if encoding bitrate is more than network bandwidth, there may be network congestion or video data loss. On the other hand, if encoding bitrate is less than network bandwidth, video quality may be degraded unnecessarily. To resolve this issue, rate-control must be employed to decide various encoding parameters to optimized transmission of video bit stream within limited bandwidth. This process ensures optimal amount of bandwidth would be utilized with maximum encoding quality [2]. Bandwidth of the channel is limited and not fixed, so that it is necessary to achieve best combination of encoding quality and bandwidth utilization. There is a trade-off between encoding quality and bandwidth utilization. Basic of rate control is depending upon target bits allocation, buffer control and rate-distortion model. There is trade-off between bit rate and fidelity. Ability of source coding system to make this trade-off well is known as rate distortion performance or coding efficiency. This system is referred as codec, comprising of coder and decoder [3].

Data compression is employed by removing redundancy. Many of the time data contains statistical redundancy and can be effectively compressed by lossless compression. Lossless compression of data doesn't compress the data by more amount. In video coding, Lossy compression achieves higher amount of compression at the expenses of loss of video quality. In a lossy compression, recovered data is not identical to the original data. In this technique, subjective redundancy is removed without significantly affecting viewer's perception of visual quality. There are two different ways of sampling series of complete frames, progressive sampling and interlaced sampling. In interlace video scanning, half of the data in a frame is sampled. A field consists of odd numbered and even numbered lines

within a video. Interlaced video sequence contains series of odd numbered fields and even numbered field, each one field has half of the information in complete frame. In interlaced video coding it is possible to send twice as many fields per second as the number of frames in an equivalent progressive sequence by keeping same data rate. This gives appearance of smoother motion. For example, The PAL video sequence, which has 50 fields per second gives smoother motion than in an equivalent progressive video sequence containing 25 frames per second [3].

Human visual system (HVS) is less sensitive to colour than luminance. In RGB colour space, equal importance is given to all three colours. So, all these colours are described with same resolution. Y:Cr:Cb is popular way of representing colour image. Y is luminance component which can be represented with higher resolution than colour component to save bits. 4:2:0 is the most popular sampling format ('YV12'), where Cr, Cb each have the horizontal and vertical resolution of Y [3].

Different video communications have different video optimum working points and over the years these points are shifted due to constraints on complexity have been eased by Moore's low [1]. Now a days implementation complexity is not major problem due growth of semiconductor technologies and advancement in computational resources. Performance of compression can be improved by taking advantage of large amount of temporal redundancy in video content, which was recognized at least as long ago as 1929 [4]

2 EXISTING STANDARDS

Two bodies are standardizing video technologies. First Video Coding Expert Group (VCEG) of International Telecommunications Union-Telecommunication Standardization Sector (ITU-T) and second is International Organization for standardization-International Electrotechnical Commission (ISO/IEC). A specific committee of ISO/IEC i.e. Moving Picture Expert Group (MPEG) standardize processing, compression, decompression, and coded representation of audio, moving picture and their combination [3]. International standards suggested by ITU-T are referred to as "Recommendation". Standardization leads to have common format across industry, vendors, video coding related software and hardware manufacturer. ITU-T recommendations are denoted with H.26x like H.265, H.264, H.263, H.262 and H.261 etc.

whereas ISO/IEC standards as denoted with MPEG-X like MPEG-7, MPEG-4, MPEG-2 and MPEG-1 etc. [6].

Evolution of various video coding standard is given in figure-1. A method of improving compression by coding only the changes in video scene is called Conditional Replenishment (CR) [7].

In first version of ITU-T Recommendation H.120, temporal redundancy reduction method was applied. In CR coding, system sends which area of a picture is to be repeated and in case if area is to be replaced, then sends updated information. Thus, CR allows a choice between skip and intra mode of representation for each area. Motion compensation prediction was missing in first version of H.120.

Hybrid codec applies two redundancy reduction techniques-using prediction and transformation. In modern hybrid block, prediction is formed from previously transmitted frames, and subtracted from current frames to form residual. A special frequency transform is applied to refinement and intra-coded region. Modern basic structure was standardized in ITU-T recommendation H.261 and continued in MPEG-1, H.262, H.263, MPEG-4 part 2 and H.264/AVC.

Some time, Fractional-sample-accurate of Motion Compensated Prediction (MCP) need more than integer precision, which need interpolation while performing MCP. Half sample accuracy MCP was considered during design of H.261 but not included due to complexity. But due to improvement in processing power and algorithms, precision of Motion Vector (MV) improved from full sample to half sample in MPEG-1, MPEG-2 and H.263, quarter sample for luma in MPEG-4, H.264/AVC and eighth sample accuracy used for chroma in H.264/AVC [1].

Motion Vector over picture boundaries standardized in H.263, which solve the problem of motion representation for sample at boundaries by replicating boundary samples for extrapolation [8]. Bipredictive MCP is averaging of two MCP signals. One prediction signal has typically been formed from a picture in the temporal future with the other formed from the past relative to the picture being predicted.

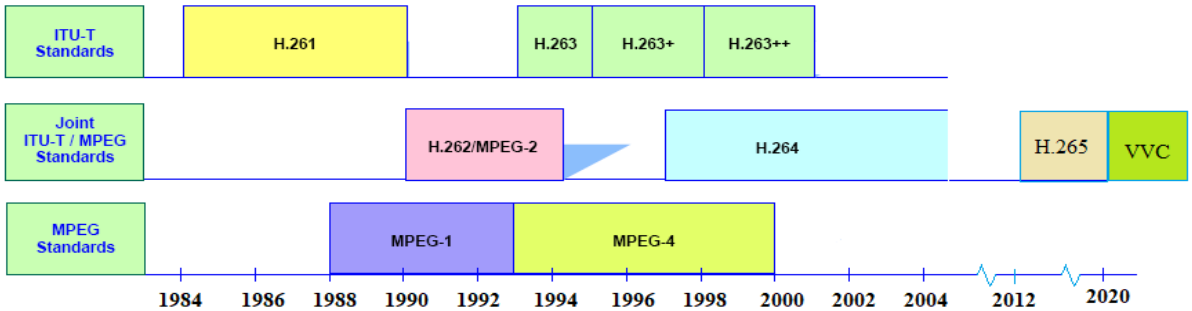

Figure 1: Development of ITU-T Recommendations and MPEG Standards

Bipredictive Motion Compensated Prediction was first put in MPEG-1 standard and continued in all other succeeding standard [9].

H.261
H.261 recommendation was adopted in 1990 as an international standard by ITU-T. This standard was designed for audio-visual service at p×64 kbps bit rate. Goal was to deliver a video in real time with typical frame rate, 15-30 fps, and minimum delay. It uses a hybrid of intra-frame and inter frame coding to exploit temporal and spatial redundancies. In intra frame coding block-based 8×8 DCT is performed and DCT coefficients are quantized. Quantized coefficients undergo entropy coding using Variable Length Huffman code to reduce bit rate. Inter frame coding involves motion compensated using motion estimation. It was originally designed to transmit video over ISDN lines. [10].

MPEG-1
MPEG-1was released in 1992 by MPEG. Primary aim was to fit compressed video into widely available CD-ROM. Apart from storage, it also deals with synchronization of audio and video information. It was different from H.261 in couple of significant ways, first, MPEG-1 uses bidirectional temporal prediction to achieve higher compression for given picture quality, second half-pixel accuracy for motion compensation.

MPEG-2
MPEG-2 standard was specified by joint technical committee of ITU-T and ISO/IEC, ratified in 1993. It was targeted for 2 Mbps or more bitrate application and video format with resolution ranging from Source Input Format (SIF) to HDTV. In this standard, audio channels are extended from 2 to 5.1. It is mostly used for multimedia data compression standard.

H.263
ITU-T releases H.263 standard in 1995, to provide significantly better quality than its predecessor. Targeted applications for this standard are multimedia on low bit rate network like Public Switched Telephone Network (PSTN), ISDN, visual telephony etc. As like previous standard, this standard use transforms coding to reduce spatial redundancy and inter-frame prediction to reduce temporal redundancy. Compared to H.261, it gives same subjective image quality at less than half bit rate. It uses half pixel accuracy for motion compensated picture.

MPEG-4
MPEG-4 standard was ratified by ISO/IEC in March-1999 as the standard for multimedia data representation and coding. It introduced object-based representation and coding methodology. In this coding scene are described in terms of foreground and background object which can be coded independently. It supports scalable coding of video objects in spatial and temporal domains. It also provides error resilience across various media.

H.264/AVC
Advanced Video Coding (AVC) is most commonly used video coding standard in industry for video recording and distribution. It is also known as MPEG-4 part 10. ISO/IEC and ITU-T formed a Joint Video Team (JVT) to develop new video coding standards. Advanced Video Coding (AVC) was ratified by JVT in 2003 which is also known as H.264. This standard became popular and adopted in many video coding applications such as iPod, Digital Video Broadcasting-Handheld (DVB-H) and Digital Multimedia Broadcasting (DMB). It supports multiple sets of features for coding algorithm that are identified to meet certain requirements of applications. These set of features are known as profile. Portable applications primarily use Base profiles while high end video coding applications such as set-of boxes, Blue-ray and High Definition DVD uses Main or High Profile at HD resolution. [11]. It is one of the most widely used coding standard such broadcast, DVD networks, packet networks and video telephony systems. This coding standard found huge success in the market. According to recent research on the market share of top online video codecs and containers, in year 2018, around 82 % of video stream on the internet are being coded with AVC/H.264 [12].

HEVC
High Efficiency Video Coding (HEVC), which is also referred as H.265 is latest video coding standard. The HEVC standard is advanced and standardized collaboratively by International Telecommunication Union-T Video Quality Experts Group (VQEG) and ISO/IEC MPEG organizations [13, Sullivan]. HEVC signifies a number of advancements in the video coding technology. Video coding layer design is based on block-based motion compensated hybrid video coding concepts, but with some changes relative to prior standards. This standard was completed in January 2013. HEVC achieve a bit rate reduction of 50% for equal perceptual video quality, and can support increased video resolution [14].
Some market participant adopted HEVC in their services and product. But market share of this coder is not very large and already appears to be saturated. The unpredicted complexity and delay in understanding the full licensing cost of implementation caused business uncertainty [15]. According to recent research on the market share of top online video codecs and containers, in year 2018, only 12 % of video stream on the internet are being coded with HEVC [12]. Even eight years after finalizing HEVC, the accurate licensing cost is unclear. Though HEVC technology is spread out to the market, licensees couldn't setup business plan due to licensing cost uncertainty.

VVC

ITU-T VCEG and ISO/IEC MPEG formed Joint Video Exploration Team (JVET). Joint Video Exploration Team (JVET) had started developing video coding techniques to improve coding efficiency, development of a Joint Exploration Model algorithm (JEM) and necessary software. In July 2020, Joint Video Exploration Team (JVET) has finalized Versatile Video Coding (VVC) standard. It increases compression capabilities by adding new tools, reaching up to approximately 50%-bit rate reduction for equivalent video quality when compared with HEVC. It is useful for emerging applications such as HDR/WCG video, 360^0 immersive video, screen contents coding etc. [16].

Average bit-rate saving for equal PSNR for entertainment applications are given below [17].

Table 1: Average bit rate for entertainment applications

Encoding	Bit-rate saving related to			
	H.264/MPEG-4 AVC HP	MPEG-4 ASP	H.263 HLP	MPEG-2/H.262 MP
HEVC MP	35.4 %	63.7%	65.1 %	70.8 %
H.264/MPEG-4 AVC HP	-	44.5%	46.6 %	55.4 %
MPEG-4 ASP	-	-	3.9%	19.7 %
H.263 HLP	-	-		16.2 %

3. CONCLUSION

In this paper, we reviewed the evolution of various video coding standards. It has been observed that, VVC is advanced video coding standard, which can compress the video more efficiently. HEVC, VVC algorithm uses high precision which are adopted due to advantage of increased computational resources. Machine learning has shown its potential in computer vision and image processing. Engineers and researcher are exploring video compression using machine learning.

However coding efficiency is not the only factor that determines choice of industry of video coding technology for services and product. Reasonable licensing terms is also important for video coding choice.

References

[1] G. J. Sullivan and T. Wiegand, "Video Compression - From Concepts to the H.264/AVC Standard," in Proceedings of the IEEE, vol. 93, no. 1, pp. 18-31, Jan. 2005, doi: 10.1109/JPROC.2004.839617.

[2] S. Xu, M. Yu, S. Fang, Z. Peng and Z. Wang, "New Rate Control Optimization Algorithm for HEVC Aiming at Discontinuous Scene," WSEAS Transactions on Computers, vol. 14, pp. 598-606, 2015.

[3] Richardson, Iain E. The H. 264 advanced video compression standard. John Wiley & Sons, 2011.

[4] R. D. Kell, "Improvements relating to electric picture transmission systems," British Patent 341 811, 1929.

[5] https://mpeg.chiariglione.org/who-we-are

[6] Sullivan, Gary J., Jill M. Boyce, Ying Chen, Jens-Rainer Ohm, C. Andrew Segall, and Anthony Vetro. "Standardized extensions of high efficiency video coding (HEVC)." IEEE Journal of selected topics in Signal Processing 7, no. 6 (2013): 1001-1016.

[7] F. W. Mounts, "A video encoding system with conditional picture element replenishment," Bell Syst. Tech. J., vol. 48, no. 7, pp. 2545–2554, Sep. 1969

[8] G. J. Sullivan and R. L. Baker, "Motion compensation for video compression using control grid interpolation," in Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP), 1991, pp. 2713–2716.

[9] T. Hidaka, "Description of the proposing algorithm and its score for moving image (A part of the proposal package)," Int. Standards Org./Int. Electrotech. Comm. (ISO/IEC) JTC 1, ISO/IEC JTC 1/SC 2/WG 8 MPEG 89/188, Oct. 1989.

[10] Akramullah, Shahriar. Digital video concepts, methods, and metrics: quality, compression, performance, and power trade-off analysis. Springer Nature, 2014.

[11] Sarwer, Mohammed Golam. "Efficient Motion Estimation and Mode Decision Algorithms for Advanced Video Coding." (2011).

[12] "Market share of top online video codecs and containers worldwide from 2016 to 2018," Statista,New York, 2019. [Online]. Available: https://www.statista.com/statistics/710673/worldwide-

[13] G. J. Sullivan, J. Ohm, W. Han and T. Wiegand, "Overview of the High Efficiency Video Coding (HEVC) Standard," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 22, no. 12, pp. 1649-1668, Dec. 2012, doi: 10.1109/TCSVT.2012.2221191.

[14] S. E. C. Osman, H. Jantan, M. T. Miskon, and W. A. K. W. Chek, "A comparative study of video coding standard performance via local area network," in International Conference on Soft Computing in Data Science. Springer, 2015, pp. 189–197.

[15] Choi, Kiho, Jianle Chen, Dmytro Rusanovskyy, Kwang-Pyo Choi, and Euee S. Jang. "An overview of the MPEG-5 essential video coding standard [standards in a nutshell]." IEEE Signal Processing Magazine 37, no. 3 (2020): 160-167.

[16] B. Bross et al., "Overview of the Versatile Video Coding (VVC) Standard and its Applications," in IEEE Transactions on Circuits and Systems for Video

Technology, vol. 31, no. 10, pp. 3736-3764, Oct. 2021, doi: 10.1109/TCSVT.2021.3101953.

[17] J. Ohm, G. J. Sullivan, H. Schwarz, T. K. Tan and T. Wiegand, "Comparison of the Coding Efficiency of Video Coding Standards—Including High Efficiency Video Coding (HEVC)," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 22, no. 12, pp. 1669-1684, Dec. 2012, doi: 10.1109/TCSVT.2012.2221192.