

## A Comprehensive Review of Traditional Video Processing

Helen Kottarathil Joy\*, Manjunath Ramachandra Kounte

School of ECE, REVA University, Bengaluru, 560064, India

### ARTICLE INFO

Article history:

Received: 08 September, 2020

Accepted: 01 November, 2020

Online: 10 November, 2020

Keywords:

Video compression

MPEG

HEVC

Deep learning

### ABSTRACT

Video and its processing are an interesting area as the increase in usage of internet videos, online streaming, CCTV, impact of internet on normal crowd increased. The need to know about video and its processing become an eminent area in research in current era. The paper tries to cover the traditional video processing, the advancement in video codec from the initial year, its origin, features, drawbacks and advancement lead to next stage. It provides an insight to need of video compression, steps involved in it, followed by overall review about video compression in various areas. The detailed explanation with reason of emergence, origin, characteristics are pointed. This information helps to add knowledge about the past and that helps to focus on the advancement and transitions that can be done to the video codecs. It summarizes the advancement in recent video processing using CNN, NN, deep learning too.

### 1. Introduction

‘Video’ the meaning has shifted far from a set of moving pictures from a traditional point of view to far extends. But the knowledge of traditional video processing is important to have a clear perspective towards the current video processing era. Video processing have a vast meaning from enhancement of various parameter in video, resolution, restoration of videos, denoising of videos, video compression etc. Each has its own impact and development in each stage of video processing development from video in a digital camera to HDTV to mobile camera to 4K,8K,10K videos. “The development of video compression started from the very started from the very start with a shadow of image compression such as Huffman coding, Golomb coding [1], arithmetic coding [2] etc. Later transform coding was introduced by encoding in spatial frequency including Fourier transform [3], Hadamard transform [4], Discreet Cosine Transform [5] etc followed by JPEG” [6]

The scope of video standardisation focus on video optimisation, allows complexity reduction for implementation and it no guarantee in terms of quality. Only decoder, bit stream and syntax of decoder is standardised. The scope if standardisation lies in the codec part of video processing. The basic knowledge about the fundamental terms in video processing helps to have strong view about any processing or standardisation done in video.

Video processing needs basic overview about the fundamentals like bitrate, display resolution, frame rate, frame type, interlacing, aspect ratio, video quality and compression techniques.

Bit rate is termed as the video data transferred in a frame of time or number of bits in the video transferred in a time duration.it is related directly to the sharpness of the frame. This is a vital thing when we consider live broadcast into consideration. The term bitrate should be handled intelligently based on the encoder and decoder. Transfer speed also is a matter to be considered here. Display resolution in video is another area to be noted in video processing. Screen resolution /display resolution was not that important part in traditional era of video processing but as time moved and screen came with various sizes colour option, video

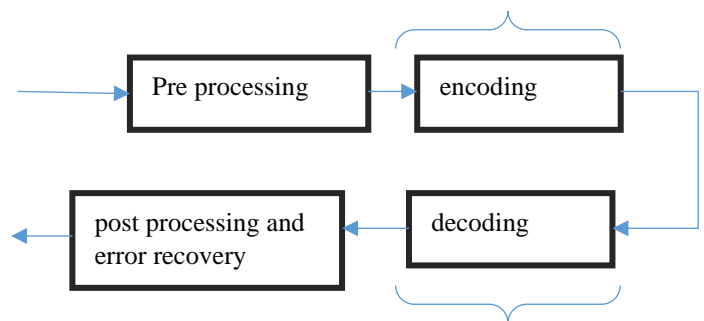


Figure 1: Scope of standardization

\*Corresponding author: Helen Kottarathil Joy, [helenjoy88@gmail.com](mailto:helenjoy88@gmail.com)

Screen resolution comes with an extension of ‘p’ and ‘i’ like 780p,1080p,1080i etc, it represents progressive /interlaced. The CRT monitors and TV followed interlaced scanning were the problem is flickering. Progressive scanning came into market as boom because the recent display screen responds fast. The refresh rate will be 60 normally to have better view effect. There are various frame resolution or frame type as in table

Table 1: Screen resolution Table

Sl.no	Resolution name	Horizontal pixels count	Vertical pixel count	Another Name
1	720p	1280	720	HD
2	1080p	1920	1080	Full HD
3	1440p	2560	1440	Quad HD
4	2160P	3840	2160	4K
5	4320P 8k	7680	4320	8K
6	4320p 10k	10240	4320	10K

Its noted that high resolution video can be viewed in low resolution screen. Video resolution doesn't matter on the screen it watching as it is taken care by the down sampling part in the decoder.

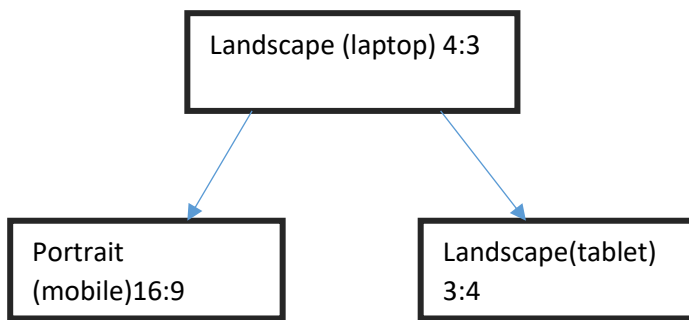


Figure 2: Aspect Ratio ideas

Aspect ratio says how the screen is with respect to the height of the display. In early ages 4:3 was common aspect ratio. In early 2010's by the boom of mobiles 16:9 become so prominent. Display orientation is another thing to be known. Size of the screen resolution should be considered properly for better display.

The paper is divided into three sections. The section 2 give brief idea about the steps in video compression, section 3 give a review of traditional video coding standards its origin, features and developmental stage towards next version. Comparison is done between various compressions followed by a overview about recent video compression with CNN and deep learning [7]. The paper ends with a conclusion with the need of learning traditional video compression for future research.

## 2. Video compression

### 2.1. Overview of video compression

Video compression is defined as a data reduction method used to encode the video. Video coding process helps that is the reduction of size of video file by making it compatible to store and need only less bandwidth for transfer. Video compression started as a succession of image compression. The compression mainly happens in a better way in video processing as the information

carried by the frames are similar in most cases. A better percentage of compression can be achieved by taking this area into consideration. Video compression is essential as storage is taken into consideration. The disk space is important as we can more videos in fact more information can be loaded. Various transforms like DCT, DST etc can also be considered to reduce file size or in other words compression.

### 2.2. The stages of video compression

The basic steps in video processing are divided into six steps starting with partitioning the picture, predicting the similarity within the frame and between the frame. Majority of compression happens in this stage [8]. The predicted frame and other information's are transformed using DCT and quantized. All these were coded by using entropy coding and send as bit streams with proper bit rate.

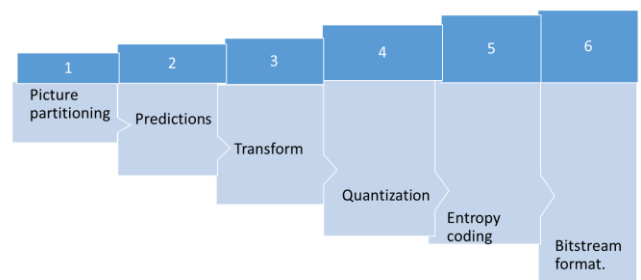


Figure 3: Basic Video Compression Steps

The input frame is transformed using DCT after picture partitioning (Splitting picture into macro blocks as apart of coding tree unit), then quantized. The quantized frame is inversed and compared with the previous frame in motion compensation and estimate is done by subtracting reconstructed frame from input frame and the residual with motion compensation. This is done by intra and inter prediction techniques basically comparing CTU within the frame and between frame to reduce redundancy and to gain maximum compression. The residual frame obtained will be 75 percentage compressed by this stage. The obtained frame 1&2 is transformed and quantized and the reconstructed frame is compared and subtracted. The process repeats and the output is encoded and framed and send in prescribed bit stream format.

## 3. Traditional video coding standards

Digital video technology covers the area of communication digital video telephony, digital TV, video storage, and a series of applications. This helps in the development of video codec in a faster way. The way of development of video codec are mainly focused on the organization ITU and ISO/IEC, joint venture of IFC and ISO are also available. ITU standards covers video codecs from H.261, H.263, H.263+, and this focus mainly in the video compression in real time communication for example video

conferencing whereas ISO/IEC mainly focus on the internal streaming, video storage etc. The bird view of video compression is shown in figure 4. Based on the purpose and application the compression features vary, upon that the segregation can be as the

one used in web or internet application, in cinemas, and for other medias. The ISO/IEC organization works on MPEG series as ITU/VCEG works on H.26x series. The ancestry of recent video compression followed this way.

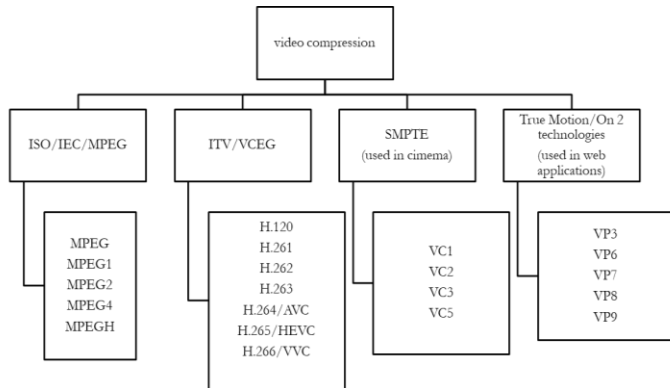


Figure 4: Video Compression Overview

Table 2: Specifications of H.262

1	Bit rate	1.5-2.0 Mbps CBR/VBR constant /variable bit rate
2	Packet/ cell loss rate (CLR)	<1 in 10 (-8)
3	BER	<1 in 10(-10)
4	Packet/ cell delay variation	<500ns

### 3.1. H.120 Video Compression Standard

The video coding standard starts with H.120 i.e. developed in the year 1984 by CCITT currently known as ITU-T. The primary application was video telephony and teleconferencing over ISDN. The bit rate available was 64kbts /sec.H.120 was not that good enough for real life applications, its spatial domain was good and temporal resolution was poor. The discussion came that encoding should have less than 1 bit as an average in a pixel. This leads to the idea of DCT block-based codec ie H.261.

### 3.2. H.261 Video Compression Standard

H.261 is the first member in the H.26x family developed in the year 1988-1990 in November by VGCE ie ITU-T9 video coding

expert groups. The focus for the development of H.261 was similar to H.120 like video conferencing in ISDN and video telephony as it couldn't satisfy it [9]. The bit rate it was focusing was multiples of 64 kbts/sec. It was used as a backward compatibility mode in H.323 and some video conferencing system still it couldn't be perfect and need refinement that leads to next video codec. But still it remains as a milestone in the history of video codecs.

### 3.3. H.262/ Moving Picture Expert Group-2

H.262/MPEG-2 was proposed in 1995 by the joint venture of VCEG and MPEG. It is mainly for storage of video like SVCD, digital video, Blue ray, broadcasting, DVD video etc. Its bit rate is suitable for only less than 1Mbps, performs bad for high bit rate. MPEG -2 is a lossy compression method with motion estimation vector, DCT quantization an encoding by Huffman coding, the technique happens in this sequence [10]. The motion vector estimation compares the frame and approximate it into similar set of video frame in a translated form with changes between the frame. This helps to reduce the greater percentage of temporal redundancy in the video frame. This is followed by DCT that helps to convert the spatial information to frequency domain and discard all high frequency information that will not affect the visual experience of human eye. Quantization is applied to the coefficient of DCT. Huffman coding shortens the code and helps in better compression but its lossy with entropy coding.MPEG is good for high quality video with a particular set of transmission limit but not that better for internet transmission as the QoS of MPEG can't be satisfied always with intend.

### 3.4. H.263 Video Compression Standard

H.263 was designed for low bitrate mainly on video conferencing. This was evolved from the pros and cons of H.261,

MPEG-1, MPEG-2 standard developed in 1996 by ITU -T9. The main application was B-ISDN video on mobile phones (3Gp), video conferencing and video telephony adopted H.324 PSTN. the bit rate followed was 33.6 Kbps. It was developed as their levels that is 1,2,3 as H.263, H.263+, H.263++ with various extra features added on it. H263+ was focusing on improved compression performance and bit stream scalability. H.263++/H.263 2000 namely MPEG 4 part 2, it helps surveillance system its simple profile formal.

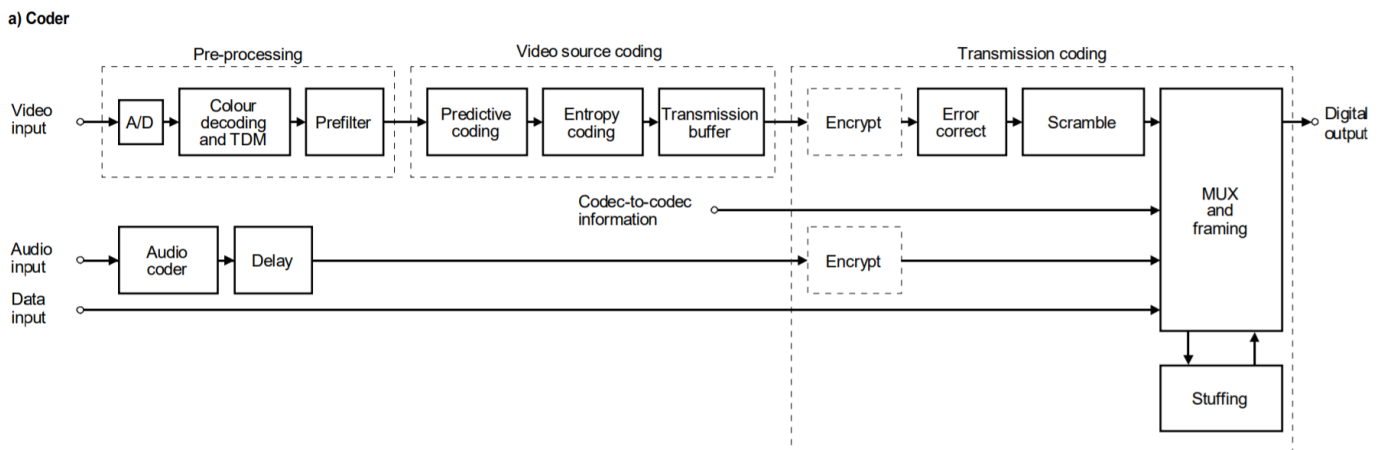


Figure 5: H.261 [7]

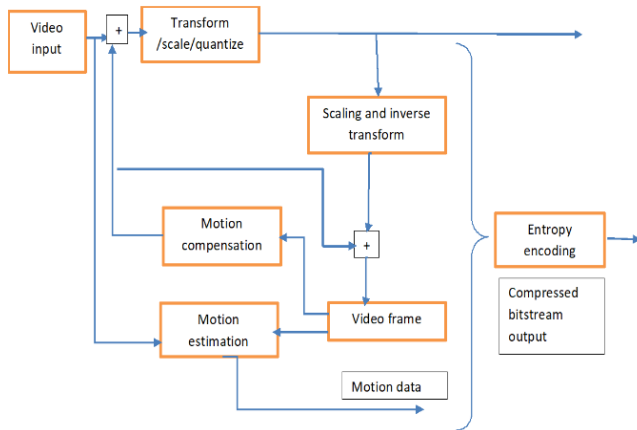


Figure 6: Hybrid video encoder [1]

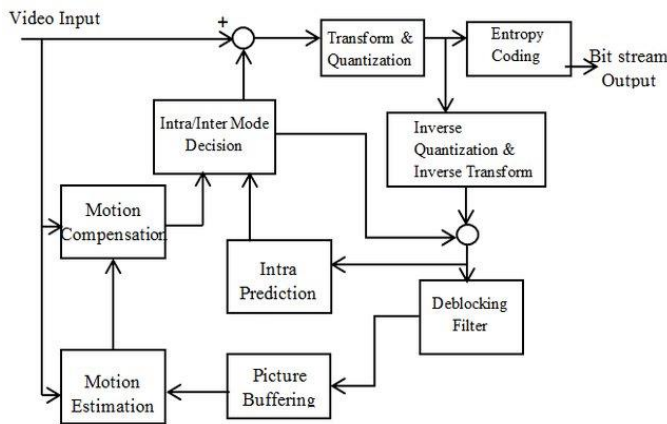


Figure 7: H.264 codec [12]

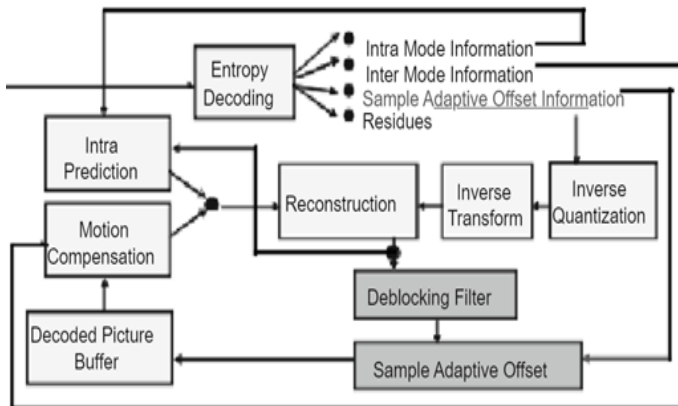


Figure 8: Video Compression with HEVC [14]

### 3.5. H.264 Video Compression Standard

H. 264/MPEG 4, AVC i.e. [10]an extended version of H.263, MPEG 4 visual released on 30<sup>th</sup> may 2003 is developed by ISO/IEC and ITU and H .264 and distribution of video content. The resolution it support is 4096X2304 that includes UHD and 4K videos. It uses variable bit rate. Advanced Video Coding /AVC is 25<sup>th</sup> version of AVC ie MPEG 4 part 10[11]. It is the best-known video coding standard for media storage like Blue ray disc and widely used in streaming in internet source that used variable bit rate. Scaling video coding is a domain feature here added with

multi view video coding, 3D-AVC and multi resolution frame compatible (MFC). The application area includes DVBC (digital video broadcast project), HD-DVD, ATSC (advanced television system connector), CCTV (closed circuit TV), DSLR etc.

### 3.6. H.265 Video Compression Standard

H.265 is developed as the successor of AVC with reduced file size, reduced bandwidth[12]. There are CTUs in HEVC than macro blocks in H.264. The HEVC have variable CTU size from 4x4 to 64x64. The main features of HEVC includes flexibility in partitioning size, flexibility in trasform block size and predicting the modes, parallel processing architecture, sophistication for prediction, signalling mode and motion vector[13]. HEVC architecture has encoder and decoder. In encoder picture partitioning is done. In the next process is the partitioned picture is predicted by intra /inter prediction and the result is subtracted from the original unit[14]. The residue is transformed by DCT and quantised in next step. At the end transformed output, predicted information, mode information, and header are entropy encoded. In decoder the counterpart does the reverse operation to deliver the picture to other end of communication[15].

### 3.7. Society of Motion Picture and Television Engineers standards

It was introduced in founded in 1916 as the Society of Motion Picture Engineers or SMPE is a global professional association, of engineers, technologists, and executives working in the media and entertainment industry. An internationally recognized standards organization, SMPTE has more than 800 Standards, Recommended Practices, and Engineering Guidelines for broadcast, filmmaking, digital cinema, audio recording, information technology (IT), and medical imaging[16-19]. In addition to development and publication of technical standards documents, SMPTE publishes the SMPTE Motion Imaging Journal, provides networking opportunities for its members, produces academic conferences and exhibitions, and performs other industry-related functions”[20].Compression Systems in SMPTE has standardized five VC standards: VC-1 to VC-5 to provide well-reviewed documentation and enhanced interoperability. The latest of these is the VC-5 standard family that provides documentation and reference software for the video compression used in GoPro systems and workflows. SMPTE also has a new project to document the Apple ProRes codec.

### 3.8. Video Processing for Internet

On 2 codec designs are popularly known as VP3 then VP7.VP6 was used in flash 8 video codec. It was used in google, skype, YouTube etc. Microsoft and Windows media player version 9 properties similar to Real video, DivX and On 2 tech. AVC improves compression quality, H.264 becomes leading standard and used by many video applications such as in iPod PlayStation Portable, as well as in TV broadcasting standards like DVB-H and DMB[21].AVS audio video standards were initialized by China, the main reason china holds the key factors in development of consumer electronics. Coding efficiency is equivalent to H.264 while computation complexity is less. It was widely used standard [22].



Table 3: Comparison Table of video codecs

Parameter	HEVC	AVC	VP9
Compression Efficiency	Twice compared to AVC	Less than HEVC	Less than HEVC
Resolution & uploading speed	Less compared to AVC	Better compared to both	More
Bandwidth required for broadcast	15mbps	32mbps	More than HEVC
Compatibility	4k TVs, VOD	3Dvideo coding.	Better in chrome, opera Firefox...
Royalty	Not open source	Not open source	Open source
Computation cost	low	high	low
Encoding quality	Good for low bit rate videos	Use lower bit rates better than previous standards	Good for high bit rate videos
FPS	Support up to 300fps	Support up to 59.94fps	Support up to 60fps for video

**4. Neural network and Deep learning-based video coding standards.**

After the development of H.265/HEVC, the blocks in the codec were explored and improved by adding neural network/ deep learning techniques [22-27] into it. The development of video codec in the following year will be mainly on making a smart video codec by incorporating the deep learning features into it[28-30]. The trial of neural network /deep learning into HEVC standard by the application of trying to make smart modules in HEVC. By using multiple features, the modules are trying to be made smart. The future of codec lies in the hands of deep learning i.e. the future is expecting a smart video codec by absorbing the basic steps in codec and improvising it by adding deep learning techniques into it. Work have been done in HEVC modules intra-prediction, inter-prediction, quantization and encoding and loop filter with neural network methods [31] that can be revised and learned separately.

As an overview the neural network for video coding can be reviewed this way, for intra prediction all the method developed for still image can be used. As we know the applications are many. For neural network based intra-prediction the effective tool considered can be motion, with the better training network it can be effectively done. Some other research happened in this area using CNN are post processing, resolution up conversion sub pixel interpolation, loop filtering, intra-coding, encoding optimization etc.

**5. Conclusion**

Video and its processing are the area that is focusing on a smart era in the near future with high quality advancement. The paper points the video processing developmental stages from initial days till now. This helps to focus on the recent research in video processing and thus paves way for new research knowing about the origin, features and properties of early codecs, earns us to work on the future of video processing. The early era of video compression started with basic coding like Huffman coding, arithmetic coding etc and transforms like Fourier transform, DCT etc followed by the hybrid video processing era with many compression techniques from Motion JPEG to H 265. The detailed study of this helps to identify that in each stage the was trying to improve the compression ratio, reduce the complexity of circuit, and with high uploading speed. The research and advancement in this area can be done by having a strong base about the existing video processing. The drawbacks of the system can be reviewed and that helps in the research to make it smarter one. The future is focusing on smart video processing and deep learning, NN, CNN are the strong pillars for intelligent video processing.

**Conflict of Interest**

The authors declare no conflict of interest

**References**

- [1] D.A. Huffman, "A Method for the Construction of Minimum-Redundancy Codes," Proceedings of the IRE, **40**(9) , 1098-1101,1952,doi:10.1109/JRPROC.1952.273898.
- [2] S.W. GOLOMB, "Run-Length Encodings," IEEE Transactions on Information Theory, **12**(3), 399-401, 1966, doi:10.1109/TIT.1966.1053907
- [3] I.H. Witten, R.M. Neal, J.G. Cleary, "Arithmetic coding for data compression," Communications of the ACM, **30**(6), 520-540, 1987, doi:10.1145/214762.214771.
- [4] H.C. Andrews , W.Pratt ,"Fourier Transform Applications", Proc. Hawaii Int. Conf. System Sciences , 677-679, 1968, doi:10.5772/2658.
- [5] W.K. Pratt, H.C. Andrews, J. Kane, "Hadamard Transform Image Coding," Proceedings of the IEEE, **57**(1), 58-68, 1969, doi:10.1109/PROC.1969.6869.
- [6] H. K. Joy and M. R. Kounte, "An Overview of Traditional and Recent Trends in Video Processing," 2019 International Conference on Smart Systems and Inventive Technology (ICSSIT), Tirunelveli, India, 848-851, 2019, doi: 10.1109/ICSSIT46314.2019.8987896.
- [7] A.S. Coyner, J.P. Campbell, M.F. Chiang, Demystifying the Jargon: The Bridge between Ophthalmology and Artificial Intelligence, Ophthalmology Retina, 2019, doi:10.1016/j.oret.2018.12.008.
- [8] A.N. Netravali, J.D. Robbins, "Motion-Compensated Television Coding: Part I," Bell System Technical Journal, **58**(7), 1703-1718,1979, doi:10.1002/j.1538-7305.1979.tb02238.x.
- [9] N. Ahmed, T. Natarajan, K.R. Rao, "Discrete Cosine Transform," IEEE Transactions on Computers, **100**(1), 90-93, 1974, doi:10.1109/T-C.1974.223784.
- [10] D.K. Kwon, M. Budagavi, V. Sze, W.S. Kim, Video compression, 2017, doi:10.1201/b12494.
- [11] T. Wiegand, G.J. Sullivan, G. Bjøntegaard, A. Luthra, "Overview of the H.264/AVC video coding standard," IEEE Transactions on Circuits and Systems for Video Technology, **13**(7), 560-576, 2003, doi:10.1109/TCSVT.2003.815165.
- [12] Wen Gao, "AVS standard - Audio Video Coding Standard Workgroup of China," 14th Annual International Conference on Wireless and Optical Communications, 2005. WOCC 2005, Newark, NJ, USA, 125–166, 2005, doi: 10.1109/WOCC.2005.1553738.
- [13] M.A. Ansari, I.U. Khan, "Performance analysis and evaluation of proposed algorithm for advance options of H.263 and H.264 video codec," in 2015 International Conference on Recent Developments in Control, Automation

- and Power Engineering, RDCAPE 2015, 371-376, 2015, doi:10.1109/RDCAPE.2015.7281427.
- [14] G.J. Sullivan, J.R. Ohm, W.J. Han, T. Wiegand, "Overview of the high efficiency video coding (HEVC) standard," *IEEE Transactions on Circuits and Systems for Video Technology*, **22**(12), 1649–1668, 2012, doi:10.1109/TCSVT.2012.2221191.
- [15] H. Lv, R. Wang, X. Xie, H. Jia, W. Gao, "A comparison of fractional-pel interpolation filters in HEVC and H.264/AVC," in *2012 IEEE Visual Communications and Image Processing, VCIP 2012*, 1-6, 2012, doi:10.1109/VCIP.2012.6410767.
- [16] K.M. Khan, J. Arshad, M.M. Khan, "Secure digital voting system based on blockchain technology," *International Journal of Electronic Government Research*, **14**(1), 53–62, 2018, doi:10.4018/IJEGR.2018010103.
- [17] V.M. Harshini, S. Danai, H.R. Usha, M.R. Kounte, "Health record management through blockchain technology," in *Proceedings of the International Conference on Trends in Electronics and Informatics, ICOEI 2019*, 1411-1415, 2019, doi:10.1109/icoei.2019.8862594.
- [18] H.K. Joy, S.L. Das, "A novel approach for biomedical web image super resolution," in *Proceedings of IEEE International Conference on Circuit, Power and Computing Technologies, ICCPCT 2013*, 876–879, 2013, doi:10.1109/ICCPCT.2013.6528858.
- [19] S.J. Kamble, M.R. Kounte, "Machine Learning Approach on Traffic Congestion Monitoring System in Internet of Vehicles," in *Procedia Computer Science*, **171**(1), 2235–2241, 2020, doi:10.1016/j.procs.2020.04.241.
- [20] S. Ma, X. Zhang, C. Jia, Z. Zhao, S. Wang, S. Wang, "Image and Video Compression with Neural Networks: A Review," *IEEE Transactions on Circuits and Systems for Video Technology*, **30**(6), 1683-1698, 2020, doi:10.1109/TCSVT.2019.2910119.
- [21] S. Matsumoto, T. Yoshihisa, T. Kawakami, Y. Ishi and Y. Teranishi, "A Distributed Video Processing System for Internet Live Broadcasting Services," *2016 19th International Conference on Network-Based Information Systems (NBIS)*, Ostrava, 2016, 311-316, doi: 10.1109/NBIS.2016.28.
- [22] M.R. Kounte, P.K. Tripathy, P. Pramod, H. Bajpai, "Implementation of Brain Machine Interface using Mind wave Sensor," in *Procedia Computer Science*, 244–252, 2020, doi:10.1016/j.procs.2020.04.026.
- [23] B. Tian, L. Li, Y. Qu, L. Yan, "Video Object Detection for Tractability with Deep Learning Method," in *Proceedings - 5th International Conference on Advanced Cloud and Big Data, CBD 2017*, 397-401, 2017, doi:10.1109/CBD.2017.75.
- [24] H. Wu, X. Zhang, B. Story, D. Rajan, "Accurate Vehicle Detection Using Multi-camera Data Fusion and Machine Learning," in *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, 3767-3771, 2019, doi:10.1109/ICASSP.2019.8683350.
- [25] M.R. Kounte, P.K. Tripathy, P. Pramod, H. Bajpai, "Analysis of Intelligent Machines using Deep learning and Natural Language Processing," in *Proceedings of the 4th International Conference on Trends in Electronics and Informatics, ICOEI 2020*, 956-960, 2020, doi:10.1109/ICOEI48184.2020.9142886.
- [26] W. Zhang, D. Zhao, L. Xu, Z. Li, W. Gong, J. Zhou, "Distributed embedded deep learning based real-time video processing," in *2016 IEEE International Conference on Systems, Man, and Cybernetics, SMC 2016 - Conference Proceedings*, 001945-001950, 2017, doi:10.1109/SMC.2016.7844524.
- [27] M.R. Kounte, P.K. Tripathy, P. Pramod, H. Bajpai, "Implementation of Brain Machine Interface using Mind wave Sensor," *Procedia Computer Science*, **171**, 244–252, 2020, doi:10.1016/j.procs.2020.04.026.
- [28] J. Lainema, F. Bossen, W.J. Han, J. Min, K. Ugur, "Intra coding of the HEVC standard," *IEEE Transactions on Circuits and Systems for Video Technology*, **22**(12), 1792-1801, 2012, doi:10.1109/TCSVT.2012.2221525.
- [29] J.L. Lin, Y.W. Chen, Y.W. Huang, S.M. Lei, "Motion vector coding in the HEVC Standard," *IEEE Journal on Selected Topics in Signal Processing*, **7**(6), 957-968, 2013, doi:10.1109/ISTSP.2013.2271975.
- [30] X. Zhang, S. Wang, Y. Zhang, W. Lin, S. Ma, W. Gao, "High-Efficiency Image Coding via Near-Optimal Filtering," *IEEE Signal Processing Letters*, 2017, doi:10.1109/LSP.2017.2732680.
- [31] L. Vinet, A. Zhedanov, "A 'missing' family of classical orthogonal polynomials," in *Journal of Physics A: Mathematical and Theoretical*, **44**(8), 085002, 2011, doi:10.1088/1751-8113/44/8/085201.