20 Years of Progress in Video Compression – from MPEG-1 to MPEG-H HEVC. General View on the Path of Video Coding Development

Damian Karwowski, Tomasz Grajek, Krzysztof Klimaszewski, Olgierd Stankiewicz, Jakub Stankowski, and Krzysztof Wegner

Poznań University of Technology
Chair of Multimedia Telecommunications and Microelectronics
Polanka 3 Street, 60-965 Poznań, Poland
dkarwow@multimedia.edu.pl

Summary. Compression of moving images has opened unprecedented opportunities of transmission and storage of digital video. Extraordinary performance of todays video codecs is a result of tens of years of work on the development of methods of data encoding. This paper is an attempt to show this history of development. It highlights the history of individual algorithms of data encoding as well as the evolution of video compression technologies as a whole. With the development of successive technologies also functionalities of codecs were evolving, which make also the topic of the paper. The paper ends the attempt of authors' forecasting about the future evolution of video compression technologies.

1 Video compression – What is it about?

The video, as we know it, no longer is a set of celluloid tapes as it used to be in the past days. Currently, practically the only form of video is a piece of digital information. Digital video is a way of storing the information about the series of rectangular array of pixels frames, representing a light in 3 frequency sub-bands: red, green and blue, changing at least 25 times per second.

The problem is that the direct representation of digital video would require an enormous number of bits. For example in the case of the nowadays television, each frame has a resolution of 1920x1080 pixels for each of the three color components, which results in more than 6 MB (megabyte: 10^6 bytes) of data. Taking into account that we have 25 such frames per second we would require 155 MB of space to store a single second of a video. And it is not the end, currently 4K, 8K or even higher resolution videos are considered. Therefore, transmission, and even storage of uncompressed video is not feasible.

In order to allow efficient video transmission, the digital signal representing the moving pictures must be compressed.

But how does one compress the video? Generally there are two approaches: to use some mathematical tricks to squeeze the data to the edge of the entropy of the data, or just to throw out less important data contained in the video.

The limits of the first approach are expressed by the entropy of the coded data. In the second approach, the question is which data are unimportant, and can be throw out. The desired way is to discard the information that is either not seen (perceived) by the viewers, or not important for the overall impression. This is the dominant approach nowadays, since it gives by the orders of magnitude better compression ratios than the lossless from definition entropy based coding.

2 What algorithms and compression technologies have been developed?

2.1 Algorithms of data encoding

In this section the main algorithms of data encoding will be recalled, which are well recognized in the field of image/video compression. From the point of view of contemporary image/video compression these algorithms should be considered as a single functional block in a codec, rather than a full compression technology.

Entropy coding

The beginning of the development of data compression methods dates back to the late 40s of the last century, when Claude E. Shannon has presented the results of his work [1], which clearly showed what are the basic limitations of data compression, e.g. how much the data can be compressed without losing information. Video signal is not an exception here.

Shannons work became the foundation for many methods of statistical coding that were developed later, which are known in the literature as entropy coding methods [2]. At least, the following methods should be mentioned here: Huffman coding [3] (year 1952), and its widely used special cases Golomb codes [5] (year 1966), exponential Golomb codes [6] (1978), LZW, arithmetic coding [4] (year 1963, but suggested earlier by Elias), and recently developed by Polish scientist ANS method [30] (year 2005).

From the beginning of development of the video compression, there were attempts to treat moving pictures as any other digital signal. Thus, one tried to encode it with the use of entropy coding only. However, it turned out very quickly that very little compression of video can be achieved in this way (fifefold more or less), due to the value of entropy of the original video data.

Prediction

Taking into account the above, additional coding tools were strongly needed, that would be an essential complement to the methods of entropy coding. First ideas of such tools boiled down to prediction of image samples based on the already transmitted ones. The goal was to transmit only the difference between prediction signal and the actual one, instead of the original image pixel.

An example of such tools were developed in 1950 in the form of Differential Pulse Code Modulation (DPCM) technique [7]. More advanced approach that was developed later is INTRA directional prediction, that is commonly used in all contemporary video codecs (like AVC and HEVC).

Even with the use of advanced INTRA prediction the efficient encoding of a video was still a serious problem. In this context, a breakthrough was in 1981, when it was developed the technique of predictive coding of a video data with the motion estimation and compensation [9]. It is commonly known as INTER coding. With this method it was possible to predict accurately the motion in a video sequence over the time, which became the basis for a very efficient compression of a video.

Transform coding

Discussed in the previous sub-section predictive coding leads to decorrelation of the data on whom this algorithm is realized. The same goal can be achieved by the use of Discrete Cosine Transformation (DCT) [8]. This is a type of a transform coding, and was developed in year 1974. In this transformation the input signal is represented with a cosinusoidal components, which in the case of image/video data can be a source of a significant reduction of a bitrate. Taking into consideration properties of human visual system it is worth to use this transformation in a combination with a lossy coding. Thus, transform coding of a video data followed by quantization of transform coefficients make an approach which is in a common use today.

2.2 Video Compression Technologies

In the previous section, examples of algorithms of data compression have been presented. Contemporary video codec must be treated as a collection of a number of such algorithms resulting in a technology (and not a single algorithm) of video compression.

Hybrid video compression

A joint application of the predictive coding (INTRA and INTER) together with the DCT-based lossy transform coding and the entropy coding of the data is commonly known in the literature as hybrid video compression. This technology of video encoding is in common use, and will be a topic of a more detailed considerations in the following part of the paper.

Wavelet image/video compression

The hybrid video coding was not the only one, which in the 80s has given a justified hopes for efficient representation of the image, also moving images. Another such method was the wavelet coding (or subband coding) that uses Discrete Wavelet Transformation (DWT) [10, 11, 12]. This method was developed especially intensive since the 80s until 2004. The performance of solutions of the wavelet compression, that were developed for still images was so astonishingly high, when compared to other available then techniques (a reflection of this performance were the capabilities of the JPEG2000 encoder [13]), that many people believed that the wavelet compression will replace the hybrid codecs in a short time. Thereby, there was increasing the pressure to repeat the success of JPEG2000 also for the purpose of a video compression, which has motivated many laboratories to work on this compression technique. Undoubtedly, a breakthrough here was the beginning of the 90s, when it was proposed the concept of a three-dimensional video coding with motion compensation (e.g. works of Ohm from 1992 – 1994 [14, 15, 16]). The success of this method caused, that the hybrid coding techniques and the wavelet techniques became for yourself a direct competition.

Situation changed in 2004 when Moving Picture Experts Group (MPEG) and Video Coding Experts Group (VCEG) looked for the best technology for scalable video coding [17]. There were number of proposals based both on the hybrid technique and wavelet coding. But at that time hybrid approach outclassed the other proposals [18]. Hybrid compression proved to be the best for compressing a video.

Parametric video compression

Video compression methods that have been cited above have become the subject of numerous applications, both in the international compression standards [19, 20, 21, 22, 23, 24], as well as commercial codecs [25, 26, 27, 28]. Over the past 20 years, other methods have been also developed, but they have not found wide practical application so far. Among these methods a special attention should be payed on algorithms of parametric coding of an image texture [31]. In the case of a texture of a respectively high degree of complexity, the data describing the texture are not sent to the decoder in general. Instead of compressing or transmitting information about individual pixels, one can try to describe the video synthetically, in words, like in the following sentence: White house built from orange bricks. Due to the bitrate of data describing parameters of the texture is incomparably smaller than a data stream of traditional encoding, one can expect its practical use in the future.

3 Milestones in history of hybrid video compression development

The history of hybrid video compression dates back to 1989 when H.261 standard has been worked out by ITU-T. Although this standard was a set of very simple coding techniques (it was DCT-based lossy coding realized in a fixed-size blocks of the image), it gave a possibility to transmit a video over ISDN networks. At the same time (more or less) the works were continued on the MPEG-1 standard of ISO/IEC, whose purpose was to compress a raw digital video under the bitrate up to 1.5 Mbps, when achieving the VHS-quality video [19].

But the standard that really revolutionized the way of video transmission and storage was MPEG-2 [20]. Developed over 20 years ago standard of ISO/IEC has proven to be vastly popular, its popularity on the market was a sign of a great success of digital television. Not only was this the beginning of the process of replacing the older analog television, but it also introduced to our homes the theater systems with a video signal of unprecedented quality. In this way, the MPEG-2 became the first sign of a certain breakthrough.

Over the years, however, expectations of the users for even higher quality of encoded video (e.g. higher video resolution, less artifacts) as well as for the larger amount of available video content (e.g. higher number of digital channels in TV) were still growing. The MPEG-2 technique offered nearly 50-fold compression of a video (while ensuring high quality of images), and that proved to be insufficient. For this reason, the hybrid compression techniques were in the following years the subject of intensive works and improvements, that resulted in developing such standards as H.263, AVS, VC-1, H.264(AVC) [21, 22, 25]. It was all the results of works that were carried out in years 1995 – 2003. Although all those standards were primarily aimed at different applications, each and every one of them introduced some new concepts, offered some improvements. Fundamental change was the introduction of a wider range of image block sizes, in which the encoder can perform data compression.

In the last 3 years, the works resulted in development of the newest, high-performance video compression technology, known as High Efficiency Video Coding (HEVC) [23]. Compared to older technologies of video compression, HEVC means more coding tools, and even higher adaptability of the size of image blocks, in which compression of data is carried out. With about 200-fold compression of a video (being a result of further improvements of the AVC technology), the new HEVC technology aims to meet the expectations of the contemporary market. Of course, this high efficiency comes at a price. The major problem is very high complexity of algorithms used in HEVC, so currently researchers look for the ways of reducing the complexity while maintaining high efficiency.

4 Is there any pattern in the chaos?

As it can be seen, the variety of solutions for the problem of video compression is broad. But surprisingly, there can be seen a pattern in the development of the new standards. Similarly as for the globally acclaimed Moore's law for the development of semiconductor devices, there is a general rule for the efficiency of the consecutive generations of video coders, expressed by Domański's curve in [29]. Beginning from the mid-nineties of the 20th century, approximately each nine years, the progress in the domain of video coding is concluded by the development of a new standard. And each new standard is approximately twice as efficient as the previous one in encoding the contents contemporary to the given video compression standard. Of course, as for the Moore's law, also for the video compression we cannot expect the trend to continue indefinitely, but it seems that we are still at the stage of continuous progress of video compression methods.

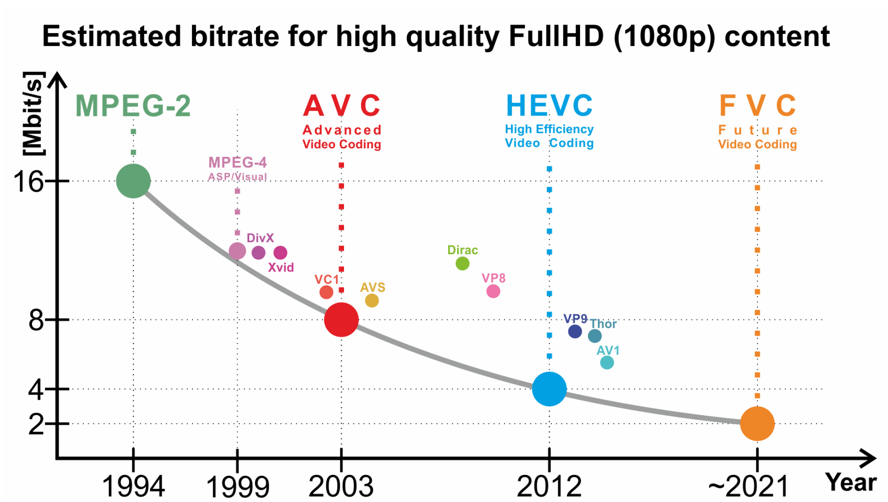

Fig. 1. Illustration of performance of successive generations of video encoders.

5 New replacing the old

It needs to be noted, though, that the newer compression methods are much more complex and much more sophisticated than the previous ones, and require much more computational power. Additionally, the proliferation of the new standard is not immediate, since the industry needs some time to adapt to the new ways of processing of the digital video. Therefore, the actual progress, as seen in the offerings of manufacturers, is not an abrupt one, but rather a steady increase of coding efficiency between the releases of the new standards.

One of the most important factors to keep in mind when considering the upgrade to the new standard is the fact, that at the beginning there is no tools to support new codec, no good practices are established and there are lots of bugs, both in software and hardware. A certain time is needed for the new standard to settle down and become usable. Trying to implement the new standard too soon often means becoming a beta tester and pioneer in the unknown.

In this connection, compression efficiency of the first industry-oriented real time video encoders is not so high. It needs some time to develop the optimal way of control of a video encoder that enables full usage of potential of the technology. General dependence of the efficiency of the reference encoder and the product encoder has been illustrated in the figure below for MPEG-2, AVC and HEVC technologies.

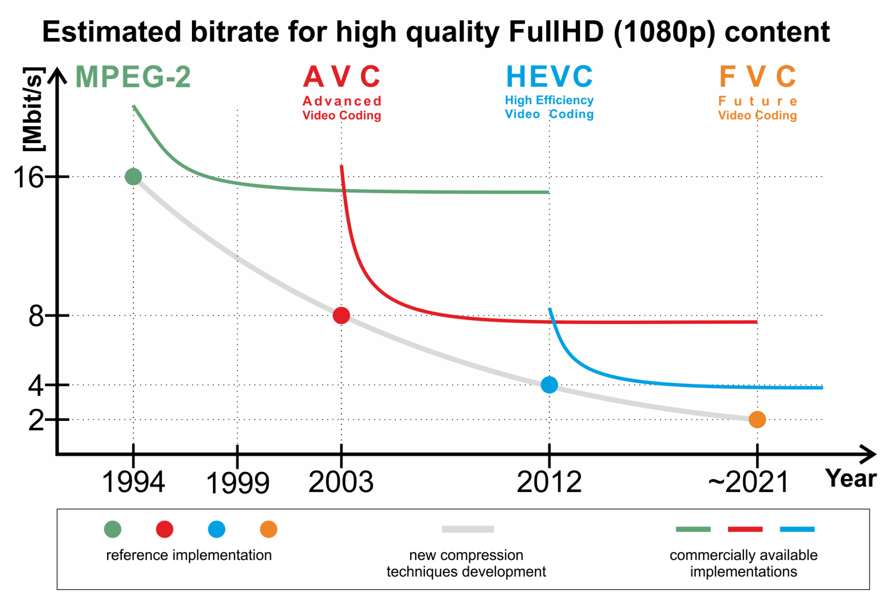

Fig. 2. Evolution of performance of commercially available encoders.

6 What was the driving force for video compression enhancements?

The most important drive for the enhancements in the process of video coding was the rising expectations of the viewers regarding the quality of the video.

Also the exponentially rising computational power of contemporary computers and more generally the hardware used for recording, processing and displaying the content contribute to the development of the new standards.

Also the progress of the science and development of the new methods and algorithms make a good foundation for development of new standards, since the companies that invent the new methods tend to capitalize on their effort by placing their developments into the coding standards.

Also the changes in the way the data is used and consumed forces the development of the new standards, in order to better meet the requirements of certain applications and to provide the new functionalities required by the given applications.

7 Evolution of functionalities

With the progress of development of compression methods, apart from efficiency, also the functionality started to matter. Not only was the efficiency started to play a role, but also other features that the codecs were expected to have. In order to maintain high coding efficiency while providing those new requested features, the coders needed to evolve by implementing a set of necessary functionalities.

One of the first significant features implemented in coders was the ability to efficiently encode interlaced sequences. In the advent of digital video, there was the necessity to compress also the video intended to use on television. In those days, television used interlaced video (each picture was divided into two fields one contained even lines of image, while the other field contained the odd lines). Specialized prediction modes were introduced to help compress those interlaced video sequences efficiently, as well as different schemes of encoding the data. In the modern television, the necessity of using interlaced video is no longer valid (the display technology changed a lot since the days of the first video coders) and the recent standard does not support interlaced video coding.

Another important feature that the codec was expected to have is the error resilience. During transmission or storage, the bitstream (i.e. compressed video data) may become corrupted. Even the change of a single bit value may render the bitstream unusable, since the entropy decoders cannot handle any errors. To prevent that, several features were introduced in the codec that help to recover from such errors, at a cost of slightly lower compression ratio. The same mechanisms make it possible to decode the bitstream from the middle such a situation is necessary for streaming of video and for digital television, since the users may start to watch the video at any time during transmission.

The next big feature introduced in video coders was so called scalability. This mechanism is intended to provide means to prepare a single bitstream that contains separate parts that make it possible to decode the video in different ways. There are different kinds of scalability:

- Time scalability when only certain frames are decoded, and the data for the other frames can be omitted
- Spatial scalability when only a lower resolution images can be decoded, while omitting data for higher resolution version of sequence
- Quality scalability when the quality of the reconstructed video depends on whether the whole bitstream is decoded or only the part of bitstream is available for decoding

The most obvious way to implement such features is to include several video streams in a single bitstream, but this leads to much larger file sizes (and thus, much wider bandwidth necessary to transmit it). But the scalability features are expected to share some parts of bitstream in order to prevent the redundancy. This means that there is a basic part of bitstream, necessary to decode any version of video, that provides the lowest quality, and there are some parts of the bitstream that improve the quality of the basic bitstream, by adding more information that enables to increase the resolution of the image, increase the quality of the images or increase the framerate. With the widespread online video streaming, this feature gains even more importance nowadays.

Another widely desired feature is the ability to compress stereoscopic video (video for left and right eye that provides the sense of depth of the scene) and even multiview images. Those functionalities were intended to be used to compress data for tv sets of all kinds that enable the 3D effect be it shutter technology, polarizer technology, optical barrier or any other 3D display technology.

The idea of transmission of stereoscopic images was soon replaced by the multiview video, where there are much more than 2 views that are encoded simultaneously. All those techniques are exploiting the fact that most of the parts of the images from different views are similar, so they present an additional opportunity for prediction of the contents of the encoded image.

The fresh extension of this idea is to encode depth maps along with a few views in order to be able to synthesize all the necessary views after decoding the video and depth maps, presumably at the decoder size. This functionality is called 3D video.

Another feature that focuses more attention due to the development of better display technologies is so called High Dynamic Range (HDR) coding. This functionality allows to encode video with more than 8 bit data for luminance and chrominances. This is a natural follow up for the development of video coders after the provision to encode very high definition video. Ultra high definition displays are able to display very real-life images, but the video data provided must not only be of sufficiently high spatial resolution, but also must have high dynamic range, so that the wider range of values for each pixel can be used. In this context it is also interesting to mention the idea of moving from a standard RGB color space of displays, to a space that allows a more lifelike images by adding a fourth basic color yellow. The development of RGBY displays has to be followed by the proper codec that could be used to provide the data for such a display.

As the cameras get cheaper and cheaper and provide even higher resolution all the time, it is also possible to imagine a system that captures the images of the whole perimeter at once those are so called 360 degrees systems. This kind of data is useful in all kinds of virtual reality applications. Of course, there is a need for a codec that is able to compress such a kind of video data.

Another new feature that the coders developed in future are expected to have are the abilities to efficiently encode computer screen content. In the past years it would be considered wasteful to record the video of a screen content, since usually it would be much more efficient (in terms of data that need to be transmitted to the receiver or stored on the disk) to simply record all the actions performed on the screen and then replicate them at the decoder side by simply replicating them.

Compared to this method, the recording of the screen content video is a much more straightforward and much easier way of storing the actions performed on-screen. There are also less compatibility issues in case of simple recording of the screen content as a video.

One of the recently added functionality is the set of modifications called Green MPEG. Those features are there to save power required to process the data at the encoder, transmission and the decoder side. This way it is, for example, possible to prolong the battery life of a mobile device that decodes the video and thus, also make the whole process of decoding a less burdening to the environment. This can be obtained by sending the additional information to the decoder to turn off or adjust the operation frequency of certain modules that will not be necessary or be used less intensively for decoding of the video or a set of frames. It is also possible to adapt the displaying process to the properties of the display, in order not to waste the energy unnecessarily, for example for the backlight.

8 Developing new encoders – change of paradigms

The earliest video compression standards (like. MPEG-2 Video) have been developed for usage in a dedicated encoding and decoding devices i.e. hardware encoders and decoders. The codec implementation cold be easily parallelized at high level (slice, frame), as well as at low level (block level) which is especially useful for hardware implementation. Briefly speaking, the goal was to develop the standard that is hardware-friendly.

Next generations of video compression technologies (H.263, AVC) have been developed in time when personal computers became powerful enough to do encoding and decoding of a video. Therefore, most compression tools that were developed in the context of H.263 and AVC took into account an efficient operation when realized in the software, and not necessary in the hardware. This led to evolution of a highly complex tools (like CABAC algorithm), which are extremely difficult to parallelize and not suitable for hardware implementations. Moreover, in abovementioned generation of compression standards there is a lack of effective parallelization techniques.

The significant turnaround took place during the development of HEVC. The coding tools complexity have been evaluated including both software and hardware implementations. The coding tools (like entropy encoding) have been substantially modified in order to allow parallel processing of data in the hardware. Moreover, a number of parallel processing possibilities have been introduced (like parallel merge, wavefront processing order, picture tiles, etc.) in order to speed up computations in the software edition of codecs.

Today we have a quite large number of different standards of video encoding. Individual standards are mutually incompatible (on the level of the syntax of encoded data stream), however most of the coding tools that are used in codecs (like motion estimation or DCT transformation) are the same. The fundamental difference is that the individual codecs may use these tools with a different input parameters, like the size of the image block in which motion estimation is carried out.

Thus, instead of implementing in a device each of the standard independently, it is better to implement a set of common functional blocks (like DCT or motion estimation), and lunch the algorithms with a given parameters, depending from the requirements of a given standard. Modern devices are built just according to this practice, which actually contain some fixed functional blocks and a general purpose core, which controls these blocks.

The architecture of devices mentioned above leads to the idea of Reconfigurable Video Coding (RVC) where encoded video bitstream contains the list and description of basic functional blocks which are necessary to decode the content.

9 What will come in upcoming years?

Of course it is impossible to predict accurately the future, nevertheless it is interesting to summarize the predicted directions in which the upcoming video compression technology will be developed. Basing on the current trends, it is anticipated that the video compression will be still constantly evolving, although without any dramatic leaps. Thanks to the progress in the area of computational hardware, e.g. processors, it will be possible to employ more and more advanced algorithms, using more and more memory, which altogether will yield higher and higher compression efficiency. As the golden fleece of computing is now usage of parallelism it is expected that the future video codecs will be designed in a way allowing for usage of multiple cores and processors with only a minor loss of efficiency.

An important factor to be considered is that the recent observations are that the pace of advancement in computational power has slowed and the progress anticipated with Moores law would soon reach saturation. Even if this will turn to be true, the speed of progress in computing should be sufficient to cater the needs of compression, because, at the same time, the image resolution enhancement is also slowing down. Even while currently 4K, 8K or even higher resolutions are considered, most of the content is displayed on 7 inch mobile devices on which Full-HD resolution is more than enough due to limitations of the humans eye. In fact, there is very little interest in representing the same field of view with higher resolutions. What is considered instead is an extension of the idea of narrow window-like vision to a wider case e.g. very wide panorama vision or even 360 vision which corresponds to representing pixels on the surface of a sphere. Such videos are already being used in applications related to light-field and virtual reality. In such, the viewers see only a part of the transmitted video, e.g. using portable viewing glasses like Oculus Rift. Therefore, partial decodability of the video stream may be of great need, because it would allow for reduction of computational costs and thus lower energy consumption. Energy consumption, connected with both battery life and power dissipation, is a fundamental concern of designers for mobile devices. The share of mobile displays in the video world is expected to rise even more and thus energy consumption might become a flagship of video compression developments. Initiatives like Green MPEG show that holistic approach to reduction of energy consumption, that will go beyond straight-forward implementation optimizations, e.g. developing energy-aware compression technologies, might be one of directions.

To summarize, it can be said the developments in video compression technology will be rather focused on new functionalities than bare compression efficiency improvements.

Acknowledgment

The research project was supported by The National Centre for Research and Development, Poland. Grant No. LIDER/023/541/L-4/12/NCBR (2013-2016).

References

1. Shannon, C.E. (1948), "A Mathematical Theory of Communication", Bell System Technical Journal, 27, pp. 379423 & 623656, July & October, 1948.
2. D. Salomon, G. Motta, Handbook of Data Compression. Springer-Verlag, 2010.
3. D. A. Huffman, A Method for the Construction of Minimum-Redundancy Codes. Proceedings of the I. R. E, pp. 1098-1101, September, 1952.
4. Abramson, N. "Information Theory and Coding", McGraw-Hill, New York, 1963.
5. S. W. Golomb, Run-Length Encoding. IEEE Transactions on Information Theory, Vol. IT-12, pp. 399-401, 1966.
6. J. Teuhola, A Compression Method for Clustered Bit-Vectors. Information Processing Letters, Vol. 7, pp. 308-311, October 1978.

7. U.S. patent 2605361, C. Chapin Cutler, "Differential Quantization of Communication Signals", filed June 29, 1950, issued July 29, 1952.
8. Ahmed, N., T. Natarajan, R. K. Rao, "Discrete Cosine Transform", IEEE Transactions on Computers, C-23:90-93, 1974.
9. Jaswant R. Jain, Anil K. Jain, "Displacement Measurement and Its Application in Interframe Image Coding", IEEE Transactions on Communications, vol. com-29, no. 12, December 1981.
10. I. Daubechies, "Orthonormal Bases of Compactly Supported Wavelets", Communications on Pure and Applied Mathematics, 41:909-996, 1988.
11. R. DeVore et al., "Image Compression Through Wavelet Transform Coding", IEEE Transactions on Information Theory, 38(2):719-746, March 1992.
12. M. Vetterli, J. Kovacevic, "Wavelets and Subband Coding", Englewood Cliffs, NJ, Prentice-Hall.
13. ISO/IEC 15444-1 and ITU-T Rec. T.800, Information Technology JPEG 2000 image coding system, 2000.
14. J.-R. Ohm, "Temporal domain subband video coding with motion compensation", Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, vol. 3, San Francisco, CA, pp. 229-232, 1992.
15. J.-R. Ohm, "Three-dimensional motion-compensated subband coding", Proceedings of International Symposium on Video Communications and Fiber Optic Services, SPIE vol. 1977, pp. 188-197, 1993.
16. J.-R. Ohm, "Three-dimensional subband coding with motion compensation", IEEE Transactions on Image Processing, vol. 3, pp. 559-571, 1994.
17. MPEG "Call for proposals on scalable video coding technology", ISO/IEC JTC1/SC29/WG11, N6193.
18. "Subjective test results for the CfP on Scalable Video Coding Technology", ISO/IEC JTC1/SC29/WG11 MPEG2004/M110737, Munich, March 2004.
19. ISO/IEC IS 11172-2 International Standard, "Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to about 1.5 Mbps", Part 2: Video, 1993.
20. ISO/IEC 13818-2 International Standard, "Generic Coding of Moving Pictures and Associated Audio ", Part 2: Video, 2nd ed., 2000.
21. ITU-T Rec. H.263, "Video Coding for Low Bitrate Communication", 2005.
22. ISO/IEC 14496-10, International Standard, "Generic Coding of Audio-Visual Objects", Part 10: Advanced Video Coding, 6th ed., 2010, take ITU-T Rec. H.264, Edition 5.0 (version 11), 2010.
23. ISO/IEC and ITU-T, "High Efficiency Video Coding (HEVC)", ISO/IEC 23008-2 (MPEG-H Part 2) / ITU-T Rec. H.265, April 2013.
24. ISO/IEC 15444-1, International Standard, "JPEG 2000 image coding system", Part 1: Core coding system, 2nd ed., 2004, also: ITU-T Rec. T.800, 2nd ed., 2002.
25. SMPTE Standard for Television: VC-1 Compressed Video Bitstream Format and Decoding Process, ANSI/SMPTE 421M 2006.
26. SMPTE Standard: VC-2 Video Compression, SMPTE 2042-1:2009.
27. SMPTE Standard: VC-3 Picture Compression and Data Stream Format, SMPTE 2019-1: 2005.
28. WebM Project, "VP9 Bitstream & Decoding Process Specification", March 2016 (v. 0.6), http://www.webmproject.org

29. M. Domaski, Approximate Video Bitrate Estimation for Television Services, ISO/IEC JTC1/SC29/WG11 MPEG 2015, M36571, Warsaw, Poland, 20-27 June 2015.
30. J. Duda, "Asymmetric numeral systems: entropy coding combining speed of Hyffman coding with compression rate of arithmetic coding", arXiv: 1311.2540.
31. P. Ndjiki-Nya, C. Stuber, T. Wiegand, "A new generic texture synthesis approach for enhanced H.264/MPEG4-AVC video coding", International Workshop on Very Low Bitrate Video (VLBV'05), Sardinia, Italy, Sep. 2005.