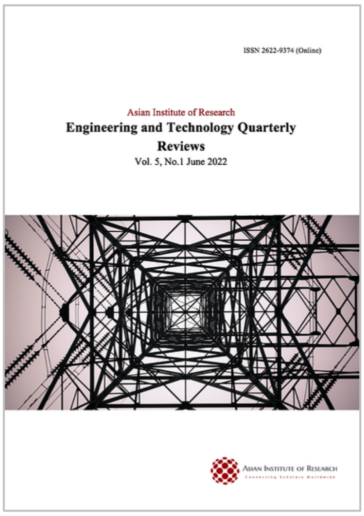

Engineering and Technology Quarterly Reviews

Jayaratne, M., Gunawardhana, L. K. P. D., & Samarathunga, U. (2022), Comparison of H.264 and H.265. In: *Engineering and Technology Quarterly Reviews*, Vol.5, No.1, 17-24.

ISSN 2622-9374

The online version of this article can be found at:
https://www.asianinstituteofresearch.org/

Published by:
The Asian Institute of Research

The Asian Institute of Research
Engineering and Technology Quarterly Reviews
Vol.5, No.1, 2022: 17-24
ISSN 2622-9374

Comparison of H.264 and H.265

Milan Jayaratne[1], Pulasthi Dhananjaya Gunawardhana[2], Uthpala Samarathunga[3]

[1,2,3] Department of Information & Communication Technology, University of Sri Jayewardenapura, Pitipana, Sri Lanka

Abstract
From the Stone Age to the present day, visual communication can be described as a form of communication that has evolved over various topics, from design to technology. As a result of globalization and the transformation of digital citizenship, services such as the Internet have come into the world. With the advent of mobile phones, computers, and television, more and more people are viewing and sharing videos, the need for a lower cost and higher quality methodology. The H.264 standard was successfully introduced in 2003 as another answer to that problem. The H.264 standard, which supported 1920 x 1080 Full HD resolutions, was no longer more practical than 4K resolutions. A decade later, with the introduction of the H265 in 2013, video usage became even more efficient. This was introduced with the combination of MPEG and ITU-T. H.265 codex mostly targets higher resolution video, such as 4k and 8k. Thru this paper, we are looking into the comparison of the technology used in H.264 and H.265. Specially H.264 and H.265 workflow, Transformation and Quantization techniques used, coding Units, Entropy Coding methods, etc. The quality of the output video is expected to be well preserved to the point where the frame concealment algorithm proposed leaves a very little negative impact on the video

Keywords: H.264, H.265, Entropy Coding, Transformation & Quantization

1. Evolution of Digital Video Industry in Brief

Man has evolved from Cave paintings from the Stone Age to the present day using various tools and techniques for his visual speech. The world's first motion picture technology, the kineograph, was created by placing a series of photographs on a cylindrical surface. With the invention of the film camera and film, cinema emerged. There were Super 8, 16mm, 35mm, and 65mm films, the most popular 35 film format.

After introducing television to the world in 1928, the world needed a more economical and convenient way of filming than the aforementioned negative films. As a solution to this, the camera and broadcast equipment manufacturers used their technologies to make DV, Mini DV, U-Matic, Beta Cam, and VHS to Create formats using different types of videotapes. Since then, video media has become more and more circulating in society with the help of digitization, which started with the release of the computer, which universities and high-tech research institutes owned, to the general market. Video technology, which had previously been used by the most luxurious elite, began to flow into the middle class.

After digitization, the media that use video in the general society can be divided into two main categories. Namely, online media and offline media. Examples include VCDs, DVDs, digital displays, video players, offline media and social media, e-learning, video conferencing, and smartphone online media. Playing (Offline Applications) or streaming (Online Applications) video with more excellent quality and the lower cost was a challenge in these wide-ranging digital video applications, and H264 and MPEG-4 compression methods were the perfect solutions.

2. History of the H.26x series

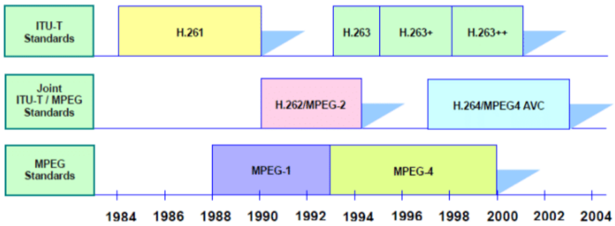

Figure 1: H.6x Family

The first compression seen in the H26x generation is the H261 compression scheme. H.261 is the International Telecommunication Union (ITU) T video coding standard approved in November 1988 (Hanzo, 2007). It was the first video codec that was useful.

H.261 was initially developed for transmission over ISDN lines, and the data rate is a multiple of 64 kbps. The encoding algorithm is designed to work with video bitrates from 40kbps to 2Mbps. H.261 supports two video frame sizes: CIF (352x288 luminance with 176x144 chroma) and QCIF (176x144 with 88x72 chroma) with 4:2:0 sampling. There is also a backwards-compatible trick to transfer still images at 704x576 luminance resolution and 352x288 chrominance resolution (Mahammad, 2017). H.263 is a video compression standard developed initially as a low bitrate compression format for video conferencing. It was developed by the ITU Video Coding Experts Group (VCEG) as part of a project that ended in 1995/1996 as part of the H.26x video coding standards suite in the ITUT domain.

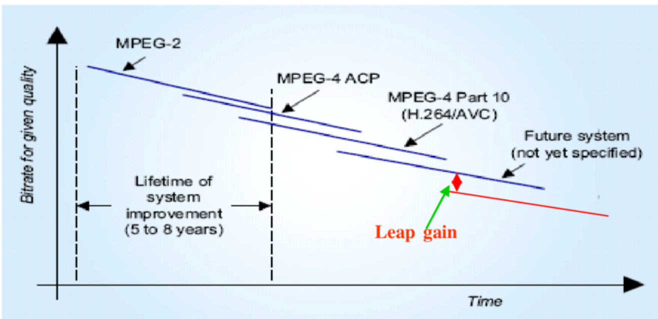

Figure 2: H.26x Family Improvement

3. Basics of H264

The H264 is arguably the most widely used video compression in the world. The H264 compression is used in broadcast, streaming, and even everyday optical discs. The H264 standard was developed by the ITU-T Video Coding Expert Group (VCEG) of the 16th Study Group with ISO / IEC JTC1 Moving Picture Experts Group (MPEG) (Recommendation ITU-T H.264, 2016). H264 Compression consists of 3 main techniques: Prediction, Transformation and Quantization, and entropy coding which is called Block-Based hybrid video coding.

3.1. Block-Based hybrid video coding

A digital video signal consists of an image sequence called a frame. This frame consists of pixels spread across two dimensions and is available in 3 main light colours: red, green, and blue. Normally this RGB pixel data is converted to Y, U, V. In this Block - Base coding system, the pixels in a complete frame are fragmented into macroblocks. There will be 16x16 size pixels as Y components and 8x8U and V components (Bull, 2014).

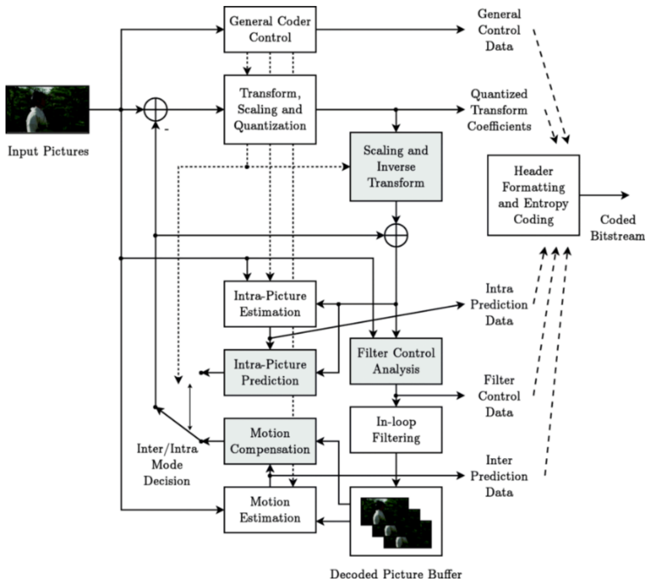

Figure 3: H.264 workflow

4. Basic Video Coding Techniques

4.1. Prediction

Prediction tries to find a reference Macroblock that is similar to the current Macroblock under processing so that, instead of the whole current Macroblock, only their (hopefully small) difference needs to be coded. Depending on where the reference Macroblock comes from, prediction is classified into inter-frame prediction and intra-frame prediction. In an inter-predict (P or B) mode, the reference Macroblock is somewhere in a frame before or after the current frame, where the current Macroblock resides. It could also be some weighted function of Macroblocks from multiple frames. In an intra-predict (I) mode, the reference Macroblock is usually calculated with mathematical functions of neighbouring pixels of the current Macroblock (Xu, 2010).

.

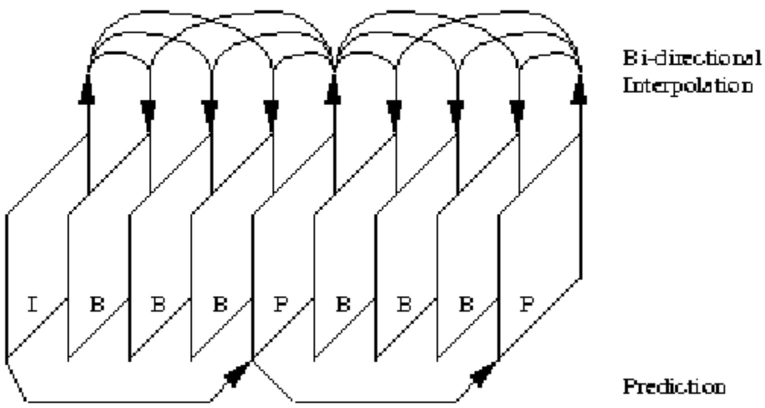

Figure 4: I, P, B frames

4.2. Compensation

According to the prediction, there are two kinds of compensation, Intra Compensation & Inter Compensation. Intra Compensation regenerates the current macroblock pixels on of 13 modes. 9 Intra 4x4 for chrominance & 4 Intra 16x16 for Luminance.

Inter Compensation (Motion Compensation) is used in a decoding path to generate the inter-frame motion predicted (estimated) pixels by using motion vectors, reference index and reference pixel from inter prediction. inter compensation also allows variable block size, multiple reference frames and quarter-pixel accurate motion vector (Wang, 2005)

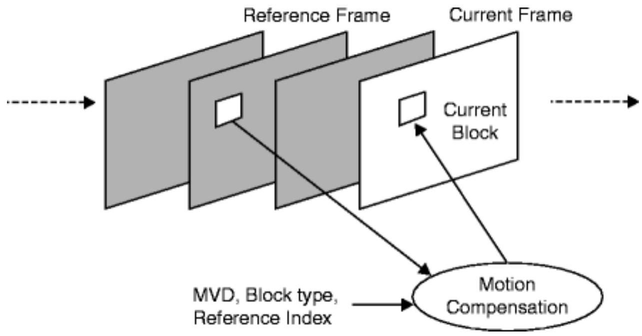

Figure 5: H.264 Motion Compansation

4.3. Transformation & Quantization

The difference between the actual and predicted data is called residual error data. Discrete Cosine Transform (DCT) is a popular block transform for image and video compression. Transforms residual data from a time-domain representation to a frequency domain representation. Because most images and videos are low-frequency data, DCT can centralize encoding information.

The main function of quantization is to reduce the transformed coefficients and the encoding information. Because the human eye is less sensitive to high-frequency image components, some video and image compression standards may use higher scaling values (quantization parameters) for high-frequency data. The h.264 standard uses the 4x4 integer DCT (Chen, 2006)

4.4. Entropy Coding

The entropy encoder is responsible for converting the syntax elements (quantized coefficients and other information like motion vectors, prediction modes, etc.) into a bitstream, and then the entropy decoder can retrieve the bitstream syntax of the elements. H.264 introduces two entropy coding methods. Context Adaptive Variable Length Coding (CAVLC) and Context Adaptive Binary Arithmetic Coding (CABAC) (Mian, 2007)

5. Basics of H265

The advent and adoption of 4K technology was a critical factor in the development of H.265. In a nutshell, 4K cameras can produce files 4x larger than regular 1080p (Full HD) files, which makes a huge difference in handling this data.

Until now, it was compressing 4K camera footage to a lower bitrate for faster streaming, and lower storage requirements often result in lower image quality than lower-compression HD footage. H.265 would theoretically eliminate this problem. However, the "theoretically" section is important because of the drawbacks associated with the H.265 implementation. H.265 or high-performance video encoding (HEVC) is sometimes the latest standard in the form of video coding and is also a promotion of H.264, which is also an extended video encoding (AVC).

The final goal of this standard is to provide the same or improved image quality. However, compressive efficiency should be improved easier to manage large data files and reduce total storage loads.

The quote varies depending on the potential savings, but several factors can affect the actual results; H.265 can reduce the bit requirements, and the relevant warehouse requirements are about 30%, and related warehouse requirements are about 30%, And related warehouses requirements are video quality. Likewise, when saving the same bit rate, close image quality is provided.

The difference is that H.265 is more aggressive in this process. In addition to changing the size from 16 x 16 pixels up to 64 x 64 or extending the checked area for pattern matching, features such as motion compensation, spatial prediction, and adaptive sampling offset (SAO) image filtering are included in compression algorithms.

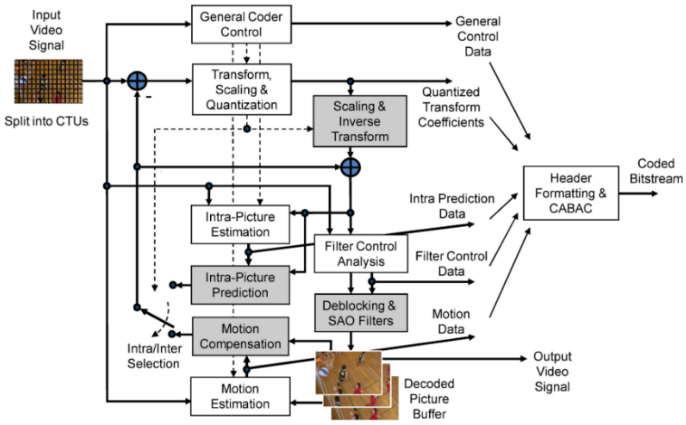

Figure 4: H.265 workflow

The H.265 system is more advanced in every aspect than the H.264 compressor system. Introducing a new hierarchical syntax representation and an Adaptive Loop Filter.

6. The Coding Tree Block

The base of the coding layer in H.264 is the macroblock, which contains a block of 16×16 luminance samples and, in the usual case, 4:2:0 chroma sampling, two 8×8 d-blocks of colour, respectively (Gary and Sullivan 2012). Response; while the same structure in HEVC is a coding tree unit (CTU), whose size is chosen by the encoder and can be larger than a traditional macroblock. The CTU consists of a luma CTB and the corresponding colour CTBs and syntax elements. The L×L size of the luma CTB can be chosen as L=16, 32, or 64 samples, with larger sizes generally allowing for better compression. HEVC then supports partitioning the CTB into smaller blocks using a quadtree-like signal and tree structure.

6.1. Coding Units (CUs) and Coding Units (CBs):

The quadtree syntax of a CTU defines the size and position of the luma and chrominance CBs. The root of the quadtree is associated with the CTU. Therefore, the size of the luminance CTB is the largest size supported for the luminance CB. The division of CTU into CB of luma and chrominance is signalled together. One luma CB and usually two chrominance CBs together with the appropriate syntax form a coding unit (CU). A CTB can contain only one CU or can be partitioned to form multiple CUs, each CU having an associated partition with a tree of prediction units (PUs) and transformation units (TUs).

6.2. Intra Prediction

Intra prediction algorithms are used to predict pixel values based on reference samples of available neighbouring blocks in the current frame. Intra Predictions contains 4 basic modes (Abramowski, 2011). There are;

Horizontal - The pixel prediction is equal to the rightmost sample in the same row in the left neighbour block.
Vertical - The pixel prediction is equal to the bottom sample in the same column in the upper neighbour block.
DC - The pixel prediction is equal to the normalized sum of the horizontal and vertical prediction.

Angular (31 Directions) - The most difficult mode requires three steps. First, the horizontal component is defined as the sample of the upper neighbour selected by the prediction angle. The scaling factor is computed based on the current row of pixels in the prediction unit. Pixel prediction equals the normalized sum of each horizontal component multiplied by each scaling factor.

6.3. Inter Prediction

The high-efficiency standard of video encoding (HEVC) improves the previous compression efficiency using a new tool and a more flexible coding structure such as Tree-Based coding block (CB / CTB – Coding Tree Blocks), prediction block (PB) and conversion block (TB). To achieve the best Rate-Distortion (RD), the HEVC reference (HM) software uses a thorough process that tests all combinations of the coding structure and test all possible combinations at the lowest RD cost. Although it provides optimal coding efficiency, this process is a significant officer that is responsible for increased calculation complexity from 9% to 502% compared to HEVC encoder, H.264 / AVC used composition.

In inter-frame prediction, the encoder initially checks the Merge / Skip (MSM) mode of CB conceptually similar to H.264 / AVC Pass mode. With MSM, you can receive information about the movement of spatially or temporary adjacent PBS and form a combined area and form joint information about the same mobile information. SKIP mode is considered a particular case of MSM. After testing MSM, the different partition modes are checked in the order shown in Figure, except for CB 8x8, which allows only the first four partition modes (MSM, 2Nx2N, 2NxN and Nx2N) (Correa, 2014).

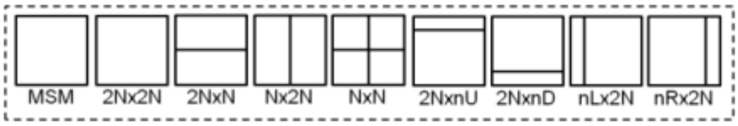

Figure 7: H265 Inter Prediction Modes

6.4. Transformation & Quantization

HEVC uses prediction error residual transform coding similar to H.264/AVC. The residual block is split into multiple square transform blocks. Supported transform block sizes: 4x4, 8x8, 16x16, and 32x32. The transformation has two main types. There are Core Transform & Alternative 4 x 4 transform. Two-dimensional transformations are calculated by applying horizontal and vertical 1D transformations direction. The elements of the underlying transformation matrix are inferred by approximating the scaled DCT basis functions, taking into account considerations such as limiting the dynamic range required for the transformation computation and optimizing maximising precision and closeness to orthogonality when matrix entries are specified as integer values.

For quantization, HEVC uses the same Quantization Parameter (QP) control scheme as in H.264/MPEG4 AVC. The QP value range is set from 0 to 51, and incrementing 6 doubles the quantization step size to map the QP values to a logarithmic approximation step size. Quantization ratio matrices are also supported. To reduce the memory required to store frequency-specific values, only quantization matrices of sizes four × 4 and 8 × 8 are used. For transformations larger than 16 × 16 and 32 × 32, an 8 × 8 scaling matrix is sent and applied by splitting the values into groups of 2 × 2 and 4 × 4 coefficients in the frequency subspace - except for those at Position DC position (zero frequency), to which discrete values are sent and applied (Sze, 2014).

6.5. Entropy Coding

Context-adaptive binary arithmetic coding (CABAC) (Sze, 2012) is the technique of entropy coding used in h.264/AVC and also in HEVC. CABAC uses three main functions: binarization, context modelling, and arithmetic coding (Sze, 2014). Binarization: HEVC uses several different binarization processes, including unary (U), truncated unary (TU), kth order ExpGolomb (EGk), and fixed length (FL). This form of binarization was also used in H.264/AVC.

Context Modeling: Context modelling provides accurate probability estimates needed to achieve high coding efficiency. Therefore, it is highly adaptable and can use different context models for different bins, and the probabilities of these context models are updated based on the values of previously encoded bins. Baskets with similar distributions often use the same context model. The context model for each bin can be chosen based on the type of syntax element, the position of the bin in the syntax element (bind X), luminance/chrominance, and neighbour information.

Arithmetic Coding: Arithmetic coding is based on recursive interval division. The range with an initial value of 0 to 1 is divided into two sub-intervals according to the probability of the bin. The encoded bits provide an offset that, when converted to a binary fraction, selects one of two subintervals representing the value of the decoded bin. After each decoded bin, the range is updated to match the selected subplot, and the interval partitioning process repeats. Ranges and offsets have limited bit precision, so whenever the range falls below a certain value, they need to be renormalized to prevent underflow. After decoding each bin, renormalization can occur.

7. Conclusion

Table 1: Summary of Comparison About H.264 & H.265

Technique	H.264	H.265
Prediction	Using 16x16 pixles Microblocks. Each consist of 16 x 16 Y (Luma Components) & 8 x 8 U,V (Chroma Components)	Using Coding Tree Unit(CTU) Instead of Microblocks and its size L X L Luma can be choosen as a L = 16,32,64
Transform & Quantization	Use 4 x 4 Integer DCT	transform matrices were derived by approximating scaled DCT basis functions
Entropy Coding	CAVLC, CABAC	CABAC

References

A. Abramowski, "Towards H.265 video coding standard," Photonics Applications in Astronomy, Communications, Industry, and High-Energy Physics Experiments 2011, DOI: 10.1117/12.905583.

"Advanced video coding for generic audiovisual services", Recommendation ITU-T H.264 [Online].2016. Available: https://www.itu.int/ITU-T/recommendations/rec.aspx?rec=13189.

C.-Y. K. Y.-L. L. Jian-Wen Chen, "Introduction to H.264 advanced video coding," Proceedings of the 2006 conference on Asia South Pacific design automation - ASP-DAC '06, 2006. ISBN: 0-7803-9451-8-/06. https://bunker2.zlibcdn.com/dtoken/901c8067304b524b0c0d610c0a3469fa

D. M. Vivienne Sze, "Entropy Coding in HEVC," Integrated Circuits and Systems, pp. 209-274, 2014. DOI: 10.1007/978-3-319-06895-4_8. https://studylib.net/doc/12703372/entropy-coding-in-hevc-please-share

D. R. Bull, "The Block-Based Hybrid Video Codec," Communicating Pictures, pp. 291-316. 2014. ISBN: 9780124059061

J. J. Y. L. Cai Mian, "An H.264 Video Encryption Algorithm Based On Entropy Coding," Third International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP 2007), 2007. DOI: 10.1109/IIH-MSP.2007.86. https://www.computer.org/csdl/proceedings-article/iih-msp/2007/29940041/12OmNBhZ4kj

J.-R. O. W.-J. H. T. W. Gary J. Sullivan, "Overview of the High-Efficiency Video Coding (HEVC) Standard," IEEE Transactions on Circuits and Systems for Video Technology, vol. 22, no. 12, pp. 1649-1668, 2012. ISSN: 1051-8215. DOI: 10.1109/TCSVT.2012.2221191

M. B. G. J. S. Vivienne Sze, High-Efficiency Video Coding (HEVC) - Algorithms and Architectures, London: Springer International Publishing Switzerland, 2014. ISBN: 978-3319068947.

M. B. Vivienne Sze, "High Throughput CABAC Entropy Coding in HEVC," IEEE Transactions on Circuits and Systems for Video Technology, vol. 22, no. 12, pp. 1778-1791, 2012. DOI:10.1109/TCSVT.2012.2221526. http://www.mit.edu/~sze/papers/sze_tcsvt_2012.pdf

P. A. L. A. L. A. d. S. C. Guilherme Correa, "Four-step algorithm for early termination in HEVC inter-frame prediction based on decision trees," 2014 IEEE Visual Communications and Image Processing Conference, 2014. DOI: 10.1109/VCIP.2014.7051505. https://ieeexplore.ieee.org/document/7051505

P. C. a. J. S. L. Hanzo, Video Compression and Communications, Wiley, 2007. ISBN: 978-0-470-51849-6.

R. W. J. W. Dawen Xu, "Prediction mode modulated data-hiding algorithm for H.264/AVC," Journal of Real-Time Image Processing, no. 4, pp. 205-214, 7. 2010. ISSN: 1861-8200. DOI: 10.1007/s11554-010-0175-4

T.-A. L. T.-M. L. a. C.-Y. L. Sheng-Zen Wang, "A New Motion Compensation Design for H.264/AVC Decoder," 2005 IEEE International Symposium on Circuits and Systems. ISBN: 0-7803-8834-8. DOI: 10.1109/ISCAS.2005.1465646. https://ieeexplore.ieee.org/document/1465646

V. M. V. Farooq Sunar Mahammad, "A Study On H.26x Family Of Video Streaming Compression Techniques," International Journal of Pure and Applied Mathematics, vol. Volume 117 No.10, no. Special, p. 63, 2017. ISSN: 4-3395.DOI: 10.12732/ijpam.v117i10.12. https://acadpubl.eu/jsi/2017-117-8-10/articles/10/12.pdf