

# CHROMA SUBSAMPLING INFLUENCE ON THE PERCEIVED VIDEO QUALITY FOR COMPRESSED SEQUENCES IN HIGH RESOLUTIONS

Miroslav UHRINA, Juraj BIENIK, Tomas MIZDOS

Department of Multimedia and Information-Communication Technologies, Faculty of Electrical Engineering, University of Zilina, Univerzitna 1, 01026, Zilina, Slovak Republic

miroslav.uhrina@fel.uniza.sk, juraj.bienik@fel.uniza.sk, tomas.mizdos@fel.uniza.sk

DOI: 10.15598/aeee.v15i4.2414

**Abstract.** This paper deals with the influence of chroma subsampling on perceived video quality measured by subjective metrics. The evaluation was done for two most used video codecs H.264/AVC and H.265/HEVC. Eight types of video sequences with Full HD and Ultra HD resolutions depending on content were tested. The experimental results showed that observers did not see the difference between unsubsampling and subsampled sequences, so using subsampled videos is preferable even 50 % of the amount of data can be saved. Also, the minimum bitrates to achieve the good and fair quality by each codec and resolution were determined.

## Keywords

*Chroma subsampling, H.264, H.265, MOS.*

## 1. Introduction

In the past, the devices designed for high-quality video recording, processing and reproduction were only domain of professional video studios. Hand in hand with the development of high-speed networks, storage media with larger space, new digital cameras and ultra-high definition video devices, the accessibility of these technologies has rapidly increased. As a result, such devices are not only used by professional but also by semi-professionals or video enthusiasts. Nowadays, 4K or 8K video recording and processing systems that support bit depth up to 16 bits and chroma subsampling up to 4:4:4 are already available.

## 2. State of the Art

Although many research activities focus on objective and subjective video quality assessment, only a few of them analyze the quality affected by chroma subsampling. In the paper [1], the efficiency of the chroma subsampling for sequences with HDR content using only objective metrics is assessed. The paper [2] presents a novel chroma subsampling strategy for compressing mosaic videos with arbitrary RGB-CFA structures in H.264/AVC and High Efficiency Video Coding (HEVC). In the paper [3], the impact of the chroma subsampling for HDR video coding is subjectively evaluated. The assessment is done only for SD, HD and Full HD resolution for four types of sequences. The paper [4] presents the measurement of the influence of two different chroma subsampling formats (4:2:2 and 4:2:0) on image quality for only MPEG-2 compression. From this search follows that publication researching the impact of chroma subsampling on the video quality using subjective assessment is missing. Therefore, the aim of this paper is to explore the influence of mentioned chroma subsampling on the video quality for the newest and most used compression standards using the selected subjective method.

## 3. Chroma Subsampling

Chroma subsampling is the process of encoding images by implementing lower resolution for chroma information than for luma information. This is due our human visual system which is less sensitive to details in colour than in luma channel. So, the video system can be optimized by devoting more bandwidth to the luma component (usually denoted Y or Y' after gamma correction), than to the colour difference components Cb

**Tab. 1:** Chrominance and luma resolution and bandwidth saving after chroma subsampling.

Chroma subsampling scheme	Chrominance resolution		Luma resolution		Bandwidth saving
	V	H	V	H	
4:4:4	Full	Full	Full	Full	0 %
4:2:2	Full	1/2	Full	Full	33 %
4:2:0	1/2	1/2	Full	Full	50 %
4:1:1	Full	1/4	Full	Full	50 %

and Cr. Table 1 represents the chrominance and luma resolution as well as the bandwidth saving. The subsampling scheme [5] and [6] is commonly expressed as a three part ratio J:a:b (e.g. 4:2:0).

The separate parts represent:

- **J**: horizontal sampling reference (width of the conceptual region), usually 4.
- **a**: number of chrominance samples (Cr, Cb) in the first row of J pixels.
- **b**: number of changes of chrominance samples (Cr, Cb) between the first and second row of J pixels. The most used subsampling schemes are 4:4:4, 4:2:2, 4:2:0 and 4:1:1.

4:4:4 - The Cb and Cr colours are sampled at the same rate as the luma (Y), thus there is no chroma subsampling (Fig. 1(a)).

4:2:2 - Both chroma components (Cb and Cr) are sampled at half the horizontal resolution of the luma (Y), so the horizontal chroma resolution is halved. This reduces the bandwidth of an uncompressed video signal by one-third compared with unsampled signal (Fig. 1(b)).

4:2:0 - Both chroma components (Cb and Cr) are sampled at half the vertical resolution of Y, so the bandwidth is halved compared to no chroma subsampling (Fig. 1(c)).

4:1:1 - Both chroma components (Cb and Cr) are sampled at one quarter the horizontal resolution, so the bandwidth is the same as by 4:2:0 subsampling scheme (Fig. 1(d)) [5] and [6].

## 4. H.264 and H.265 Compression Standards

Compression is one of the most important parts of the video transmission system that has an impact on the video quality. During the past two decades, many video compression standards have been developed. The most common ones are based on MPEG platform.

Although H.264/AVC codec, developed in 2003 is still one of the most used compression standards. It has been designed for a wide range of applications, ranging from video for mobile phones through web applications to TV broadcasting (HDTV). H.264/AVC also defines profiles and levels. There are only three profiles currently defined - Baseline, Main, Extended [7].

The High Efficiency Video Coding known as HEVC/H.265 was developed in January 2013 by a partnership known as the Joint Collaborative Team on Video Coding (JCT-VC) which arose by the cooperation of the ITU-T Video Coding Experts Group (VCEG) and the ISO/IEC Moving Picture Experts Group (MPEG) standardization organizations. It is the newest coding standard from the MPEG family codecs. It contains many improvements which make it more effective than the previous standards [8].

## 5. Subjective Video Quality Assessment

Generally, the video quality evaluation can be divided into two groups - the subjective and the objective one.

The subjective quality assessment can be defined as a perceived quality. It relies on people - observers - who watch the video sequences and rate the quality. It is the most reliable and fundamental way how to determine the video quality called Quality of Experience (QoE). The biggest advantage of this type of measurement is the accuracy of the results; the most major drawback is the duration of the assessment - it is very time consuming. From the aspect the number of stimuli (reference and impaired sequences), the subjective methods can be divided to [9] and [10]:

- single stimulus methods (e.g. ACR - Absolute Category Rating, SSCQE - Single Stimulus Continuous Quality Evaluation),
- double-stimulus methods (e.g. DSIS - Double Stimulus Impairment Scale, DSCQS - Double Stimulus Continuous Quality Scale).

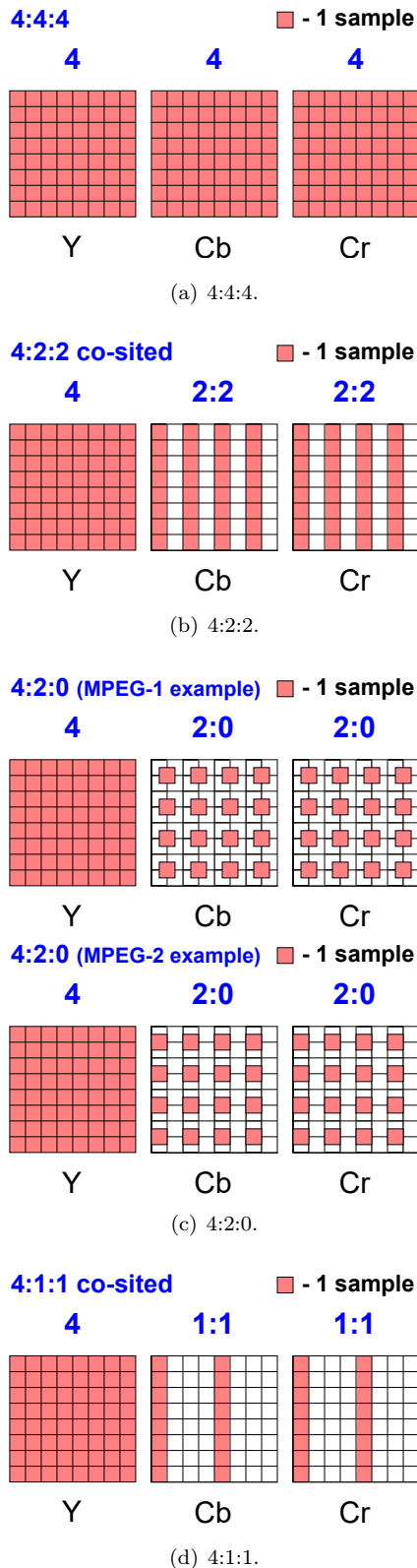


Fig. 1: Chroma subsampling modes.

The subjective methods can be also divided depending on the fact when the quality measurement is performed:

- methods in which the quality is evaluated after the presentation (e.g., DSIS, ACR),
- methods in which quality assessment takes place during the test sequence - continuous quality assessment (e.g. SSCQE, SDSCE).

All necessary information, as well as all metrics, are defined in ITU-R BT.500-13 [9] and ITU-P. 910 recommendations [10]. In our research, the Absolute Category Rating method was chosen and used.

### 5.1. The Absolute Category Rating Method (ACR)

The ACR method, also called the Single Stimulus method (SS), is a type of measurement when only the impaired (the test) sequence is shown to the observer (Fig. 2), so the viewer does not know which quality is the reference sequence.



Fig. 2: Presentation structure of the ACR method.

The assessor is asked to rate the quality of the test sequence based on the level of the quality he has in his opinion for it after watching it. The five-level grading scale is used: 5 for excellent, 4 for good, 3 for fair, 2 for poor, 1 for bad [9].

## 6. Measurement

In the paragraphs below, the whole procedure of the measurement is described. It consists of three parts - the test sequences description, the coding process and the subjective assessment.

### 6.1. Test Sequences

In our measurement eight test video sequences depending on content (Fig. 3) were used. All sequences are the part of the database [11] and were downloaded in uncompressed format (\*.yuv). Basic parameters of these sequences are shown in Tab. 2. Regarding to [10], the spatial and temporal information indicates the type of content and is directly related to compression efficiency. Due to this reason, the Spatial (SI) and the Temporal Information (TI) of all sequences using the Mitsu tool [12] were computed and are shown in SI and TI diagram (Fig. 4). Below a short characteristic of each sequence is written.





(a) Bund Nightscape.



(b) Campfire Party.



(c) Construction Field.



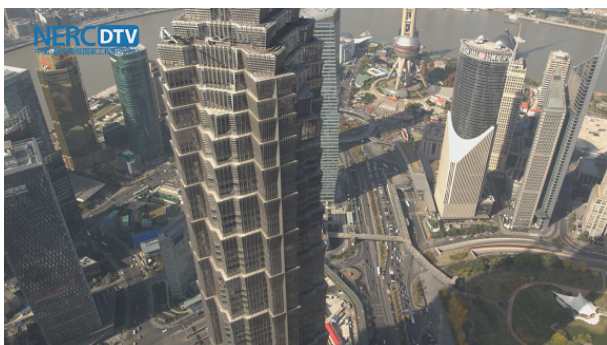
(d) Fountains.



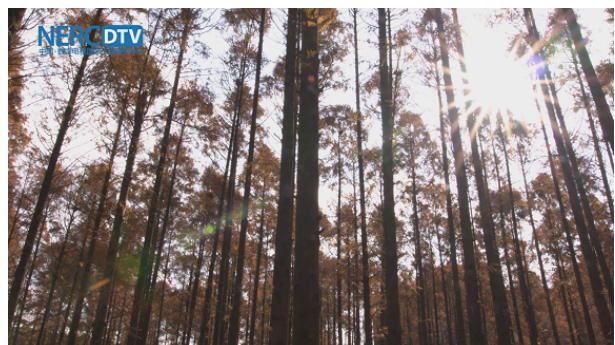
(e) Marathon.



(f) Runners.



(g) Tall Buildings.



(h) Wood.

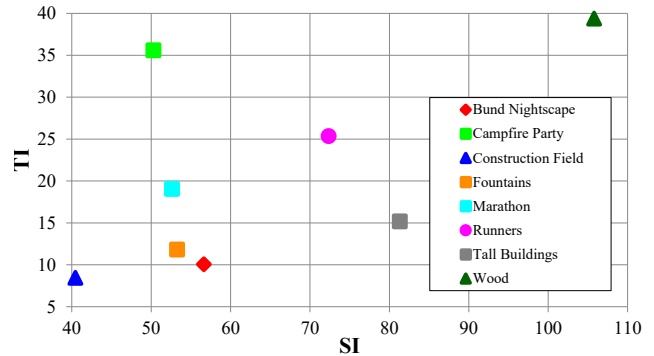
Fig. 3: Test sequences.

**Tab. 2:** Parameters of the test sequences.

Resolution	Chroma subsampling	Bit depth	Aspect Ratio	Framerate (fps)	Length (seconds)
3840×2160 (UHD)	4:4:4	10 bit per channel	16:9	30	10

### 1) Characteristics of Used Test Sequences

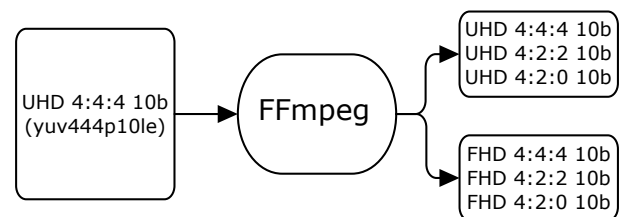
- *Bund Nightscape*: time lapsed night city scene captured by static camera. The static parts are represented by buildings and the horizon, walking people and driving cars are the only dynamic objects in the scene (Fig. 3(a)).
- *Campfire Party*: night scene of the fire in the front of the image and group of people in the background. The flaming bonfire is changing quickly (the fast change of temporal and luminance information). Group of people in the background is moving slowly. At the end of the sequence, the camera zooms on the group of people (Fig. 3(b)).
- *Construction Field*: slow-motion scene of the building site with the static background. The only dynamic objects are excavator and walking workers. The scene is captured by the static camera (Fig. 3(c)).
- *Fountains*: view of the city fountain. The squirting water in the foreground (a lot of edges in the image). The background is static and consists of trees and buildings. The camera is static, the scene with a minimum of motion (Fig. 3(d)).
- *Marathon*: marathon competition captured from the static point of view. Runners represent moving objects; the background is static street (Fig. 3(e)).
- *Runners*: relatively dynamic scene of running competition, but unlike to "marathon" scene, there are fewer runners. The camera is static and the runners are running closer to the camera. The camera is angled to the side (higher spatial information) (Fig. 3(f)).
- *Tall buildings*: bird's eye view of the modern city. The static objects are the skyscrapers, river and city infrastructure. The slow-motion objects are cars. The camera is panning slowly (Fig. 3(g)).
- *Wood*: shot of the trees in the forest. The camera is moving from the left to the right side and the speed of moving is slightly increasing. Relatively high value of the temporal and spatial information (Fig. 3(h)).

**Fig. 4:** SI and TI diagram.

## 6.2. Coding Process

The coding process was done in two steps:

First, all downloaded video test sequences were chroma subsampled from 4:4:4 format to 4:2:2 and 4:2:0 formats, as well as the resolution from UHD to FHD, was changed. This was done using FFmpeg tool [13], version 3.2.4. build with gcc version 6.3.0. In this step, six test sequences in the uncompressed format were created. This process was done for all types of test sequences (Fig. 5).

**Fig. 5:** Process of preparing the test sequences - chroma subsampling and resolution changing.

Second, all created test sequences were encoded to both H.264/AVC and H.265/HEVC compression standards. The target bitrates were set exactly to 1, 3, 5, 10 and 15 Mbps by setting the bitrate options in the FFmpeg tool (see Tab. 3). The GoP was set to the half of the framerate, i.e.  $M = 3$ ,  $N = 15$ . For coding, again the FFmpeg tool, version 3.2.4 was used [13]. The command line settings of this tool for both compression standards are shown in Tab. 3. In this step for the assessment 240 sequences for each resolution were created.



Tab. 3: Command line settings of the FFmpeg tool.

Options settings	FFMPEG command line settings
Input options	-i Input,Test Sequence.yuv, -video_resolution 3840×2160, -pix_fmt yuv444p10le, -framerate 30
Codec option	-vcodec libx264 (libx265)
GoP options	-keyint=15, -minkeyint=15, -bframes=3, -b-adapt=1
Bitrate options	-bitrate=bitrate in Mb·s <sup>-1</sup> , -vbv-maxrate=max bitrate in Mb·s <sup>-1</sup> , -vbv-buFSIZE= max buFSIZE in Mb·s <sup>-1</sup>
Output options	Output,Test Sequence.mp4

### 6.3. Subjective Assessment

Finally, the video quality of all sequences using ACR method was evaluated. For the assessment, the quality ratings of observers were used. For the measurement, the home environment according to [9] was chosen and the Samsung LE40C750 display type was used. The complete process of the measurement and evaluation is shown in Fig. 6.

The data of observers who watched and evaluated the quality is shown in Tab. 4.

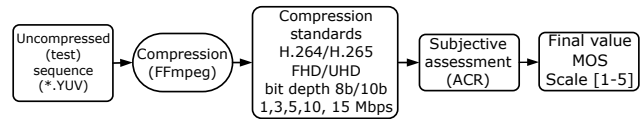


Fig. 6: Complete process of coding and evaluating the quality of the test sequences.

Tab. 4: Data of the observers.

Resolution	Number of men	Number of women	Average age
FHD	22	8	24
UHD	17	13	22

## 7. Experimental Results

The next figures show the impact of bitrate on the video quality (MOS scale) measured by the ACR method. By all graphs, the confidence interval was also computed. All results in the plots are done only for most used 4:2:0 type of chroma subsampling.

Figure 7 shows the impact comparison of the used type of scene. Each curve represents each of used test sequences. In this figure, four graphs are inset depending on the codec and the resolution.

According to the graphs we can say that the sequences with low SI and TI values as the "Bund Nighscape" and the "Construction Field" reach the

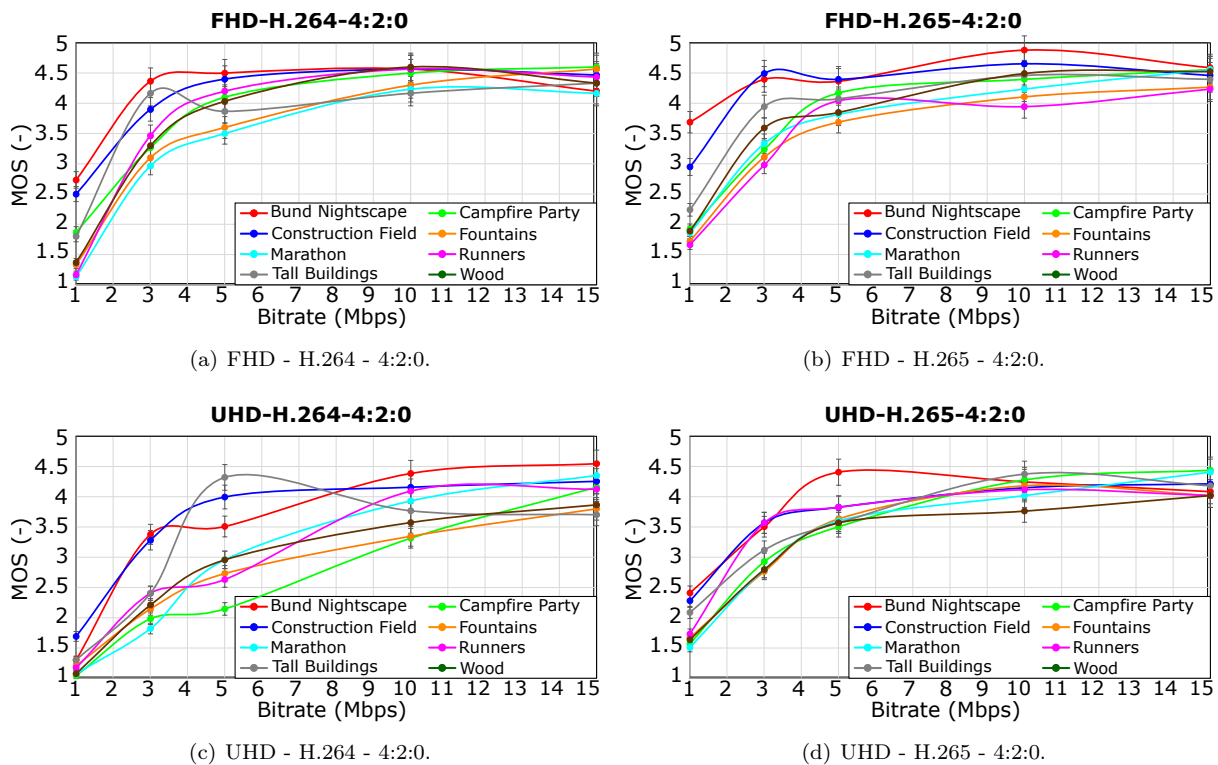


Fig. 7: Scene comparison.

best MOS score by low bitrates. The difference is more visible in Full HD resolution than in Ultra HD resolution. Vice versa, the sequences with higher SI and TI values as "Marathon" or "Runner" reach lower MOS score by low bitrates. With increasing bitrate, the quality rises too and approach the quality of sequences with low SI-TI values. A particular case is the "Tall Buildings" sequence which is in the plot situated between two upper mentioned groups of sequences.

In the next step, we calculated the average MOS score from all used test sequences for each codec and resolution and plotted it in Fig. 8. So, Fig. 8 shows the impact comparison of used codec and resolution.

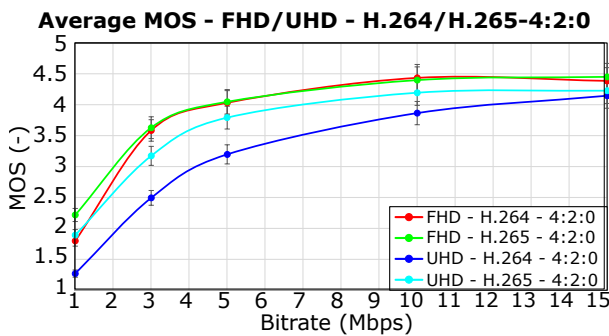


Fig. 8: Codec and resolution comparison - average MOS score.

We also computed the differences between H.265 and H.264 codec by both resolutions. The results are expressed in MOS score and shown in Tab. 5. The performance of both codecs is compared using the bitrate saving characteristic which is represented by Tab. 6.

From Fig. 8, Tab. 5 and Tab. 6, we can see that

Tab. 5: Codec comparison expressed in the MOS score.

Bitrate (Mb·s <sup>-1</sup> )	FHD	UHD
	H.265 - H.264 Difference in MOS scale	H.265 - H.264 Difference in MOS scale
1	0.43	0.64
3	0.05	0.70
5	0.02	0.61
10	-0.04	0.34
15	0.07	0.08

Tab. 6: Codec performance comparison expressed by the bitrate saving characteristic.

Bitrate (Mb·s <sup>-1</sup> )	FHD	UHD
	H.265 - H.264 Bitrate saving	H.265 - H.264 Bitrate saving
1	19.81 %	34.77 %
3	1.27 %	22.12 %
5	0.41 %	16.10 %
10	-0.85 %	8.05 %
15	1.50 %	1.97 %

H.265 codec achieves better quality than H.264 codec. This fact is generally known and we assumed it. We

can also state that the quality difference between used codecs is bigger in Ultra HD resolution than in Full HD resolution, so the quality between these two codecs is for the observers better recognized in Ultra HD resolution than in Full HD resolution. In Full HD resolution, the difference is visible only in bitrates between 1 and 3 Mb·s<sup>-1</sup> (0.43 MOS score; 42.92 %). Despite it, in Ultra HD resolution, the difference is bigger (for 1 Mb·s<sup>-1</sup> 63.75 %, for 3 Mb·s<sup>-1</sup> 69.58 %, for 5 Mb·s<sup>-1</sup> 60.83 %) and the curve representing the H.264 codec approaches the H.265 curve even in higher bitrates (for 10 Mb·s<sup>-1</sup> 33.75 %, for 15 Mb·s<sup>-1</sup> 8.33 %). Also, regarding the same Fig. 8, Tab. 5 and Tab. 6, the quality of the Ultra HD sequences was rated by the observers with worse MOS score than the quality of the Full HD sequences. It is due the fact that the Ultra HD video sequences contain more information, i.e. more data than the Full HD sequences, so the quality degradation caused by the compression reflects and is more visible in Ultra HD resolution than in Full HD resolution.

From Fig. 8, we can also determine the minimum bitrates to which should be the video sequence coded to achieve good (4) or fair (3) quality. These quality thresholds are based on MOS scale of used ACR method. In Tab. 7, the mentioned minimum bitrates are written.

Tab. 7: Minimum bitrates to achieve good and fair quality.

MOS scale	FHD (Mb·s <sup>-1</sup> )		UHD (Mb·s <sup>-1</sup> )	
	H.264	H.265	H.264	H.265
Good (4)	5	5	12	7
Fair (3)	2.2	2	4.25	1.75

Table 7 follows that to achieve the good quality, the video sequence should be coded to minimum 5 Mb·s<sup>-1</sup> by both codecs (H.264 and H.265) for Full HD resolution and to 12 Mb·s<sup>-1</sup> by H.264 codec and to 7 Mb·s<sup>-1</sup> by H.265 codec for Ultra HD resolution. To reach the fair quality, the video sequence should be coded to minimum 2.2 Mb·s<sup>-1</sup> by H.264 and to 2 Mb·s<sup>-1</sup> by H.265 codec for Full HD resolution and to 4.25 Mb·s<sup>-1</sup> by H.264 and to 1.75 Mb·s<sup>-1</sup> by H.265 codec for Ultra HD resolution.

The last Fig. 9 shows the impact of chroma subsampling on the video quality; four graphs are inset depending on the codec and the resolution. The MOS values are shown in Tab. 8.

According to the graphs and Tab. 8, we can declare that the difference between unsubsampled and subsampled videos is not recognized - the observers did not see the difference between the sequences coded to 4:4:4 subsampling format and sequences coded to 4:2:2 or 4:2:0 subsampling formats. It follows that using subsampled videos is preferable - people cannot see the difference and even 50 % of the amount of data can be saved (see paragraph 3 - Tab. 1).

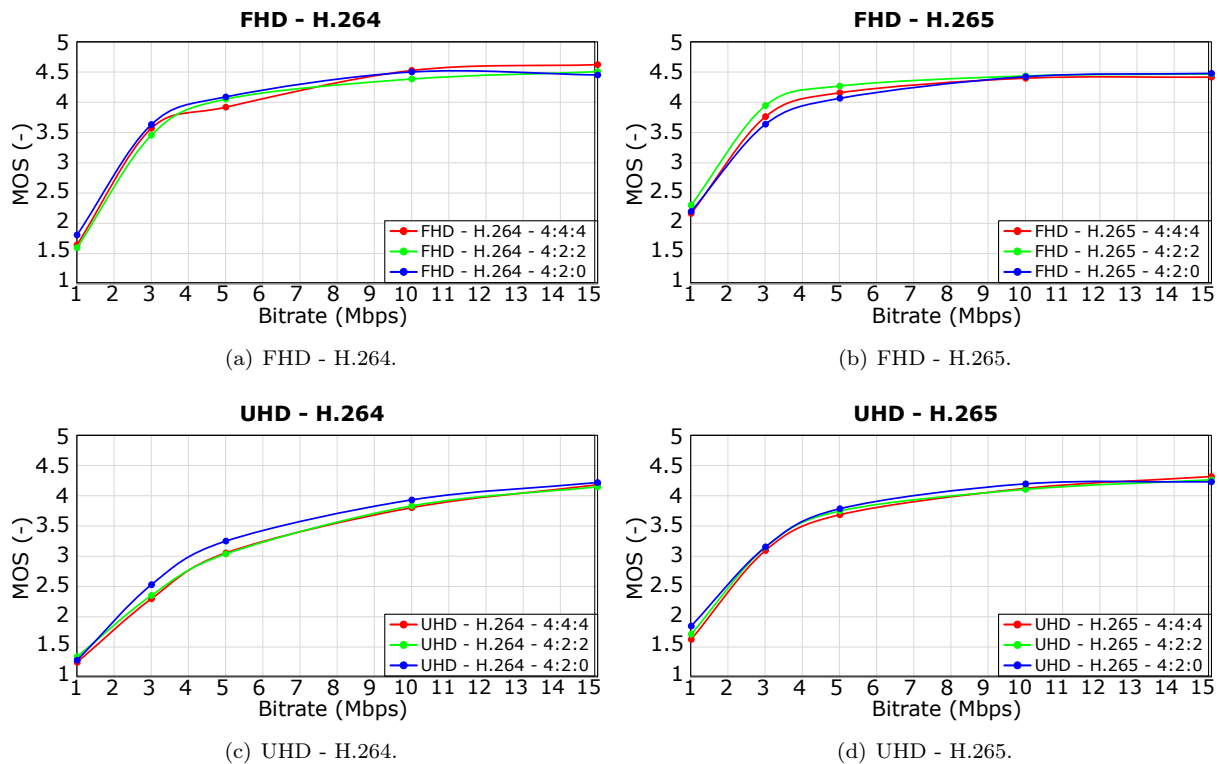


Fig. 9: Chroma subsampling comparison.

Tab. 8: MOS values of used subsampling formats for both codecs and resolutions.

FHD						
Bitrate (Mb·s <sup>-1</sup> )	H.264			H.265		
	4:4:4	4:2:2	4:2:0	4:4:4	4:2:2	4:2:0
1	1.58	1.53	1.74	2.13	2.27	2.17
3	3.50	3.39	3.57	3.74	3.92	3.61
5	3.85	3.99	4.03	4.13	4.25	4.04
10	4.46	4.32	4.44	4.38	4.41	4.40
15	4.56	4.44	4.39	4.39	4.45	4.45
UHD						
Bitrate (Mb·s <sup>-1</sup> )	H.264			H.265		
	4:4:4	4:2:2	4:2:0	4:4:4	4:2:2	4:2:0
1	1.16	1.25	1.20	1.61	1.70	1.83
3	2.22	2.27	2.45	3.09	3.14	3.15
5	2.98	2.96	3.17	3.68	3.74	3.78
10	3.73	3.75	3.85	4.12	4.10	4.19
15	4.10	4.07	4.14	4.31	4.26	4.23

## 8. Conclusion

This paper dealt with the influence of chroma subsampling on perceived video quality measured by subjective metrics. The evaluation was done for two common video codecs H.264/AVC and H.265/HEVC. Eight types of video sequences with Full HD and Ultra HD resolutions depending on content were tested. The experimental results showed that the difference between unsubsampled and subsampled videos is unrecognizable - the observers did not see the difference between the coded sequences using 4:4:4 subsampling method and coded sequences using 4:2:2 or 4:2:0 subsampling

methods. This suggests that using subsampled videos is preferable - even 50 % of the amount of data can be saved. Also, the codec performance comparison expressed in MOS score as well as by the bitrate saving characteristic was done. At the end, according to the tests, the minimum bitrates to achieve the good and fair quality by each codec and resolution were determined.

## Acknowledgment

This work was supported by Centre of excellence for system and services of intelligent transport II, IMTS 2622120050 supported by the Research & Development Operational Programme funded by the ERDF.

## References

- [1] BOITARD, R., M. T. POUZARD and P. NAISOPOULOS. Evaluation of chroma subsampling for high dynamic range video compression. In: *23rd IEEE International Conference on Electronics, Circuits and Systems*. Monte Carlo: IEEE, 2016, pp. 696–699. ISBN 978-150906113-6. DOI: 10.1109/ICECS.2016.7841297.



- [2] LIN, C.-H., K.-L. CHUNG and C.-W. YU. Novel Chroma Subsampling Strategy Based on Mathematical Optimization for Compressing Mosaic Videos with Arbitrary RGB Color Filter Arrays in H.264/AVC and HEVC. *IEEE Transactions on Circuits and Systems for Video Technology*. 2016, vol. 26, iss. 9, pp. 1722–1733. ISSN 1051-8215. DOI: 10.1109/TCSVT.2015.2472118.
- [3] THOMA, H., M. DE-FRUTOS-LOPEZ and J. AUER. Chroma subsampling for HDR video with improved subjective quality. In: *Picture Coding Symposium*. San Jose: IEEE, 2013, pp. 345–348. ISBN 978-147990294-1. DOI: 10.1109/PCS.2013.6737754.
- [4] DUMIC, E., M. MUSTRA, S. GRGIC and G. GVOZDEN. Image quality of 4:2:2 and 4:2:0 chroma subsampling formats. In: *ELMAR-2009 - 51st International Symposium ELMAR-2009*. Zadar: IEEE, 2009, pp. 19–24. ISBN 978-953704410-7.
- [5] WINKLER, S., C. J. VAN DEN BRANDEN LAMBRECHT and M. KUNT. Vision and Video: Models and Applications. In: *Vision Models and Applications to Image and Video Processing*. Boston: Springer, 2001, pp. 201–229. ISBN 978-1-4757-3411-9.
- [6] Definition of: chroma subsampling. In: *PC-mag* [online]. 2017. Available at: <http://www.pcmag.com/encyclopedia/term/57460/chroma-subsampling>.
- [7] ITU-T H.264. *Advanced video coding for generic audiovisual services*. Geneva: ITU, 2004.
- [8] ITU-T H.265. *High efficiency video coding*. Geneva: ITU, 2017.
- [9] ITU-R BT.500-13. *Methodology for the subjective assessment of the quality of television pictures*. Geneva: ITU, 2012.
- [10] ITU-T P.910. *Subjective video quality assessment methods for multimedia applications*. Geneva: ITU, 2009.
- [11] Test Sequence. In: *Test Sequence* [online]. 2013. Available at: <http://medialab.sjtu.edu.cn/web4k/index.html>.
- [12] Video Quality Indicators. In: *Video Quality* [online]. 2017. Available at: <http://vq.kt.agh.edu.pl/metrics.html>.
- [13] FFmpeg tool. In: *FFmpeg* [online]. 2017. Available at: <https://www.ffmpeg.org/>.

## About Authors

**Miroslav UHRINA** was born in 1984 in Zilina, Slovakia. He received his M.Sc. and Ph.D. degrees in Telecommunications at the Department of Telecommunications and Multimedia, Faculty of Electrical Engineering, at the University of Zilina in 2008 and 2012, respectively. Nowadays he is an assistant professor at the same department. His research interests include audio and video compression, video quality assessment, TV broadcasting and IP networks.

**Juraj BIENIK** was born in 1987 in Zilina, Slovakia. He received his M.Sc. in Telecommunications at the Department of Telecommunications and Multimedia, Faculty of Electrical Engineering, at the University of Zilina in 2012. Nowadays he is a Ph.D. student at the same department. His research interests include audio and video signal processing, functionality and optimisation of networks and video quality assessment.

**Tomas MIZDOS** was born in 1993 in Poprad, Slovakia. He is received his M.Sc. in Telecommunications at the Department of Multimedia and Information-Communication Technologies, Faculty of Electrical Engineering, at the University of Zilina in 2017. Currently he is the Ph.D. student at the same department. His research interest focuses on audio and video signal processing and video quality assessment.