Testing HEVC model HM-16.15 on objective and subjective way

Zoran M. Miličević, Jovan G. Mihajlović and Zoran S. Bojković

***Abstract*—**This paper seeks to provide performance analysis for High Efficient Video Coding (HEVC) HM-16.15 reference coder for same resolution test sequences and picture structure. Two test sequences (resolutions) i.e., PeopleonStreet (2560x1600) and Traffic (2560x1600) are considered. The HEVC coder is tested for the fixed Quantization Parameter (QP) value in lowdelay and randomaccess configurations, with Main profile when B pictures are processed in IBBB format, respectively. Comparisons were performed with respect to the change of Signal-to-Noise Ratio (SNR), the change of data bit-rate and the change of encoding time saving, respectively. Simulation results have shown small differences in SNR values and obvious differences for bit rate and encoding time values. Beside objective results, subjective video assessments for all tested sequences are presented, as well as present of intra and inter prediction in total amount encoder compression values. At the end, motion vectors and coded block structures are analyzed.

***Keywords*—**bit-rate, encoding time saving, HEVC test model (HM), signal-to-noise ratio (SNR).

I. INTRODUCTION

IN the family of video coding standards, HEVC has got the potential to replace/supplement all the existing standards , i.e. MPEG, as well as H.26x series including H.264/AVC, starting from the point that with the continuous growth of video resolution, the standard H.264/AVC cannot give expected compression results. For example, in order to improve the quality of the reconstructed video and to improve the compression ratio, video coding standard HEVC was invoked. Of course, a limited increase in computer complexity compared to the H.264/AVC standard was obtained [1]. As a result, we have high consumption increase power, dissipation and limited processing times. It should be pointed out that the encoding computational complexity of HEVC depends on the Test model (HM) configuration used [2] - [4]. When going to compare HM software test models, it can be done though three fundamental parameters: signal-to-noise ratio, bit rate and time saving, while processing test sequences in different resolutions.

Zoran M. Miličević is with the Department of Telecommunication and IT GS of the SAF, Raska 2, 11000 Belgrade, Serbia (phone: 381-64-1191125; e-mail: zmilicko@yahoo.com).

Jovan G. Mihajlović is with the Department of Telecommunication and IT GS of the SAF, Raska 2, 11000 Belgrade, Serbia (phone: 381-69-4405063; e-mail: mihajlovicjovan@yahoo.com).

Zoran S. Bojković is with the University of Belgrade, 11000 Belgrade, Serbia (e-mail: z.bojkovic@yahoo.com).

II. HEVC APPROACH

First version of the HEVC H.265 standard was received by the ITU-T in 2013, although recent progress in many various extensions was carried out. The time H.264/AVC was under standardization, a few devices supported high-definition (HD) videos. As a successor of H.264/AVC, HEVC was designed with the goal to satisfy emerging demands of high-quality video services, for example HD TV [5]. The basic processing unit in HEVC is longest coding unit (LCU) which is no overlapped squared block. Each coding unit (CU) can be divided into four partitions.

Today, researchers are exploring the way how to reduce the HEVC encoders complexity. The focus has been in reducing the motion estimation (ME) complexity, because ME occupies 77%-81% of HEVC encoders implementation. Performance comparison of HEVC with older standards such as H.264/AVC, MPEG-4 Part 2 Visual, H.262/MPEG Video, H.263 and also with image coding standards such as JPEG, JPEG 2000, was carried out. [6]. Also, researchers are exploring transcoding between HEVC and other standards such as MPEG-2 and H.264. Further extensions of HEVC are scalable video coding (SVC), 3D video multiview video coding and range extensions which include screen content coding (SCC), bit depth larger than 10 bits and color sampling of 4:2:2 and 4:4:4. In general, SCC refers to computer generated object, both images and videos requiring lossless coding. Some of these extensions have been finalized by the end of 2014, while time frame for SCC was late 2016. Iguchi et al. have already developed a hardware encoder for super hi-vision (SHV) i.e., with HDTV at 7680x4320 pixel resolution. Also real time hardware implementation of HEVC encoder for HD video has been done. HEVC standard supports both progressive and interlaced scan methods.

Also, HEVC standard include a set of profiles which define the configuration and tools that must be used to provide interoperability between various implementations for a particular case. Additionally defining a set of profiles, HEVC standard as well define a set of conformance levels. They define the highest bitrate, spatial and temporal resolutions that may be used. Decoder that accepts the standard must fit the current level and all lower levels.

In addition to the introducing of HEVC standard coding and transform tools, the principle used to partition pictures vary intensely from the 16x16 pixel macroblock employed in the

H.264 family of codecs. A quadree coding structure has been adopted in which a coding unit has a highest size of 64x64 pixels, that can be beyond sub-devided in to smaller units 8x8 pixels size. Quadtree splitting depth depends on comparing the distortion cost rate of subdividing to that of not subdividing at each quadtree depth level.

HEVC standard quadtree coding structure example is shown in Fig. 1.

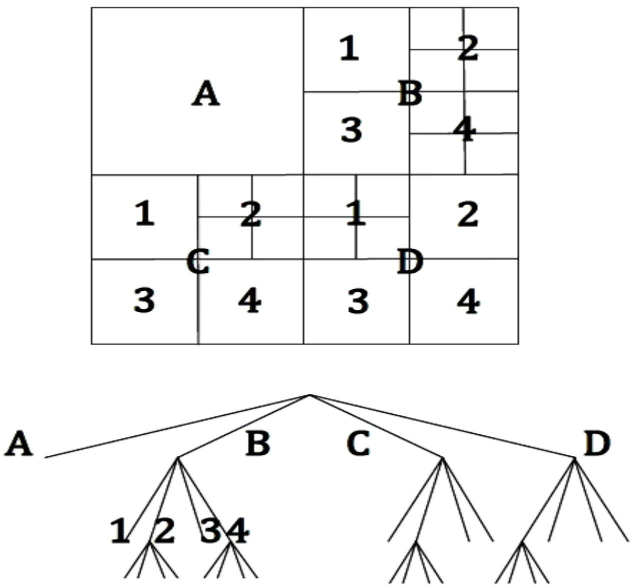


Fig. 1. The quadtree coding structure of HEVC standard.

Logical data unit known as network abstraction layer or NAL unit ensure a storage and transport format for the encoded picture data created by VCL layer.

As in all prior ITU-T and ISO/IEC JTC 1 video coding standards since H.261, the HEVC design follows the classic block-based hybrid video coding approach. The basic source-coding algorithm is a hybrid of interpicture prediction to exploit temporal statistical dependences, intrapicture prediction to exploit spatial statistical dependeces, and transform coding of the prediction residual signals to further exploits spatial statistical dependences. There is no single coding element in the HEVC design that provides the majority of its significant improvement in compression efficiency in relation to prior video coding standards. It is, rather, a plurality of smaller improvements that add up to the significant gain [5].

III. Challenges For Real-Time Applications

The increased encoding complexity represents one of the very important challenges for real-time applications. The quad-tree structure for coding unit with different sizes and a large number of prediction modes is one of the reasons for encoding complexity of HEVC. Thus, one of the challenges for real-time applications is to develop a test mode decision method for reducing computational complexity for HEVC. Secondly, several different methods have been investigated recently, aiming of computational complexity reduction and scaling of HEVC software implementations. Thus, maintaining the encoding time for frame or group of pictures (GOPs) below an adjustable upper bound is still an open research issue.

IV. Simulations Results

Simulation results represent the continuation of our experimental work on performance evaluation for various versions of HM software test model in different conditions.

We evaluated the performance of the HEVC model HM-16.15 [7], when encoder_lowdelay_main and encoder_randomaccess_main configurations were used. The system platform was the Intel(R) Core(TM) i3-2328M Processor of speed 2.2 GHz, 6 GB RAM, and Microsoft Windows 7 Professional. The HEVC software configurations were as follows: Main profile, value of Level: 5.0, I pictures, hierarchical B pictures, period of I-pictures: only first (of both configurations), Hadamard transform was used, Maximum coding unit depth was 4, MV (Motion Vectors) search range was 64, SAO (Sample Adaptive Offset), AMP (Asymmetric Motion Partitions) and RDOQ (Rate-Distortion-Optimized Quantization) were enabled, GOP (Group of Pictures) length 16 in encoder_randomaccess_main and 4 in encoder_lowdelay_main for IBBB format were used. The QP (Quantization Parameter) used was 32.

All processed configurations are adopted to Main profile.

Experiments were carried out on the tested sequences with fixed quantization parameter value QP=32. We chose QP=32 as value of the QP, because it is approximately average value in reference software setup configuration.

For the experiments two different test sequences are selected. The selected test sequences are in same resolution and frame rates. We used the first 50 frames of test sequences Traffic and PeopleonStreet. The both test sequences in resolution 2560x1600 pixels belongs to class A. All the test videos are in YUV 4:2:0 format and progressive. Details about the test sequences and sequence classes that are used for the comparisons in the paper are summarized in [8].

Also, the SNR values of luma (Y) component of pictures are used. We measured SNR only for Y because human visual system is more sensitive to luma then to chroma components of pictures.

Comparisons with the case of exhaustive search were performed with respect to the change of Signal to Noise Ratio (SNR), the change of data bit-rate (Bit-rate), and the change of encoding time saving (Time), respectively.

Table 1 shows the performance and comparison of the reference codecs for hierarchical B pictures processing in the IBBB format in lowdelay and randomaccess configurations for QP=32, respectively, based on our simulation results.

When both tested sequences are compared there are small differences in SNR values (5% for lowdelay configuration and the little bit over 6 % for randomacces configuration denoted by „-") for luma component of picture in the test model HM-16.15.

TABLE 1. PERFORMANCE AND COMPARISON OF THE REFERENCE CODECS BASED ON SIMULATION RESULTS.

Test sequences (resolution)	**Profile** *Main*	**SNR-Y (dB)**	**Bit-rate (kbps)**	**Time saving (sec)**
People on Street (2560x1600)	Lowdelay	34,45	8881,21	22306,44
	Randomaccess	34,09	7501,66	14371,62
Traffic (2560x1600)	Lowdelay	36,17	2482,89	15279,06
	Randomaccess	36,22	1959,81	9648,65
Results	Lowdelay	-5,01	72,04	31,50
	Randomaccess	-6,24	73,88	32,86

From the bit rate point of view, for both test sequences there are obvious differences in values for luma component of picture in both tested configuration. The bits rate of test sequence PeopleOnStreet is increased 72% in the lowdelay configuration, while in the randomaccess configurations it is approximately 74% in comparison with Traffic test sequence.
Finally, for PeopleonStreet test sequence the encoding time saving is increased 31,50% for lowdelay and 32,86% for randomaccess configuration when that test sequence is compared with Traffic test sequence.

(a)

(b)

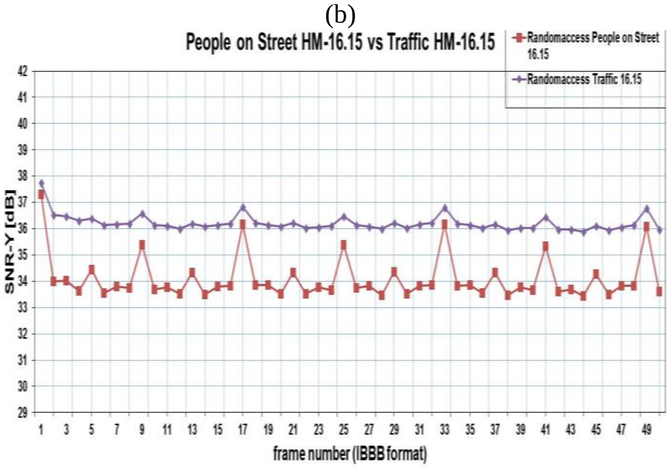
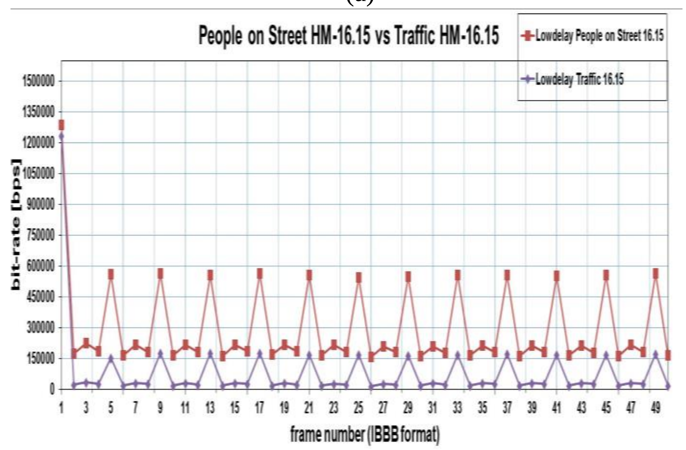


Fig. 2. The SNR curves for PeopleOnStreet vs. Traffic test sequences when different configurations in IBBB picture format are processed in HM-16.15 reference software.

SNR curves for PeopleOnStreet vs Traffic test sequences are depicted in Fig. 2 ((a) and (b)). The SNR-YUV is plotted as a function of the frame number for both tested configurations. For both processed test sequences, SNR shows differences on objective way.

Bit-rate savings curves for both typical tested sequences are depicted in Fig. 3 ((a) and (b)) for the lowdelay and randomaccess configurations, respectively. There exist the bit rate differences between both the HEVC tested configurations, as well as bit-rate trends, as it has been shown in Table 1.

(a)



(b)

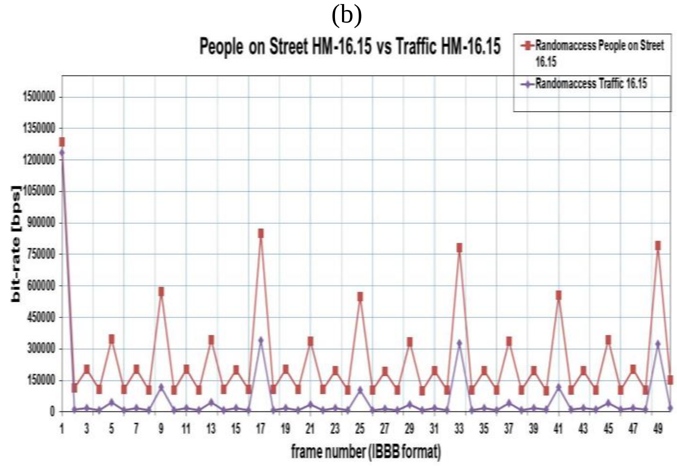
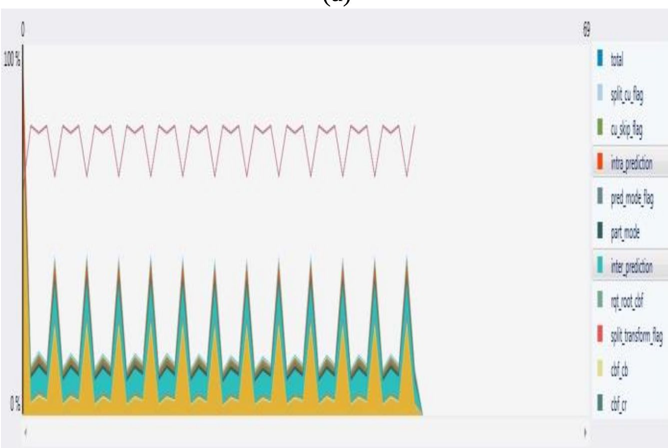


Fig. 3. The Bit-rate curves for PeopleOnStreet vs. Traffic test sequences when different configurations (lowdelay and randomaccess) in IBBB picture format are processed in HM-16.15 reference software.

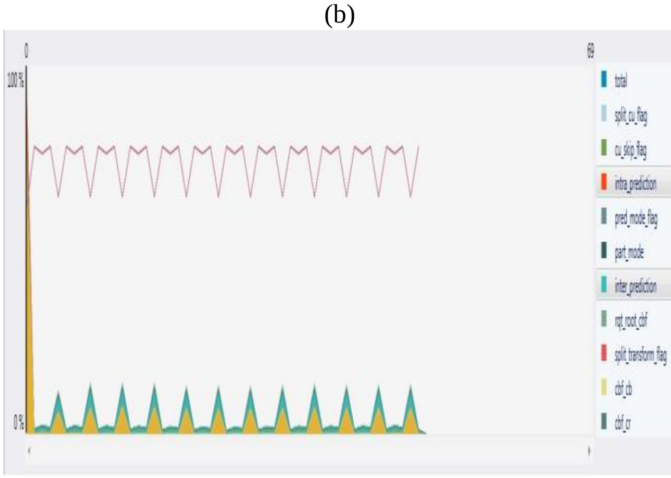
Fig. 4 ((a), (b), (c) and (d)) shows HEVC HM-16.15 total amount encoder compression values when Elecard_hevc_analyzer is used. In Fig. 4 a) there is highlighted percent of intra and inter prediction for PeopleonStreet test sequences in lowdelay configuration (red and light blue marker in the shadow area). From total amount of values for intra prediction share is 5,40%, while for inter prediction share is 29,08%. Also, in Fig. 4 b) there is highlighted percent of intra and inter prediction for Traffic test sequences in lowdelay configuration. In this case, intra prediction share is 4,38%, while inter prediction share is 22,92%. Next, in Fig. 4 c) there is highlighted percent of intra and inter prediction for

PeopleonStreet test sequences in randomaccess configuration. The intra prediction share is 7,07%, while inter prediction share is 31,79%. Finally, in Fig. 4 d) there is highlighted percent of intra and inter prediction for Traffic test sequences in randomaccess configuration. In this case, intra prediction share is 6,54%, while inter prediction share is 21,11%. Output results shows that intra prediction fluctuates from 4% to 7%, while inter prediction fluctuates from 21% up to 32% in sum of total values depend of test sequence and tested configuration.

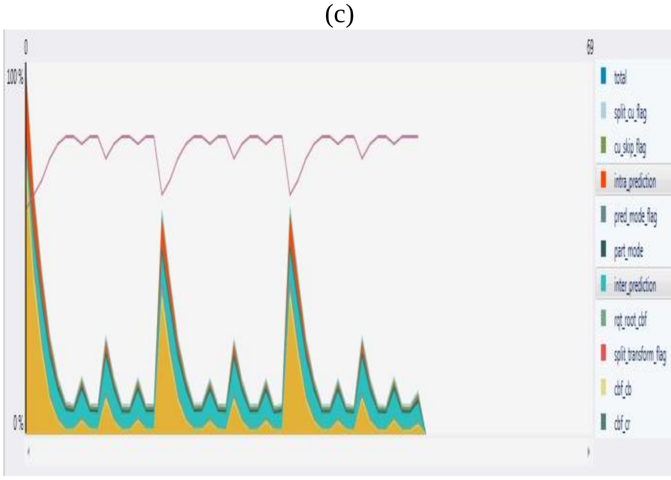
(a)



(b)



(c)



(d)

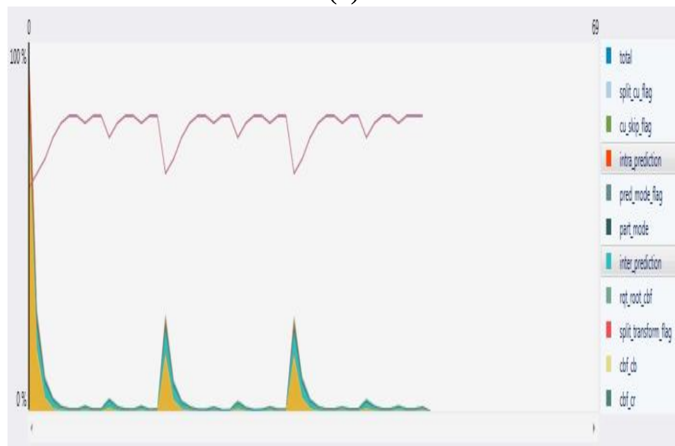


Fig. 4. HEVC intra and inter prediction percent in total amount encoder compression values for PeopleOnStreet and Traffic test sequences processed in HM-16.15 reference software.

Besides objective analysis of the HEVC encoders for two different resolution test sequences, subjective video quality is analyzed, too.

Fig. 5 ((a), (b), (c) and (d)) shows HEVC HM-16.15 in both tested sequences, configurations and IBBB picture format for subjective video assessment, respectively. All tested sequences are processed by YUV player, respectively. Subjective assessment results clearly indicate that there are small differences in term of SNR in Fig. 2 in accordance with results in Table 1.

(a)

People on Street - lowdelay HM-16.15

(b)

Traffic - lowdelay HM-16.15

(c)

(d)

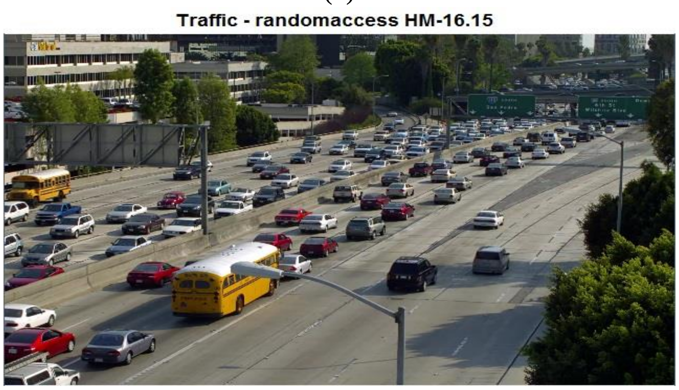
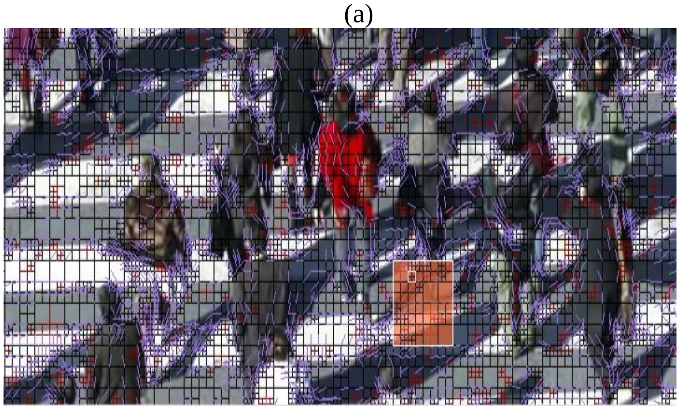


Fig. 5. HEVC subjective video assessment for PeopleOnStreet and Traffic test sequence when different configurations in IBBB picture formats are processed in HM-16.15 reference software.

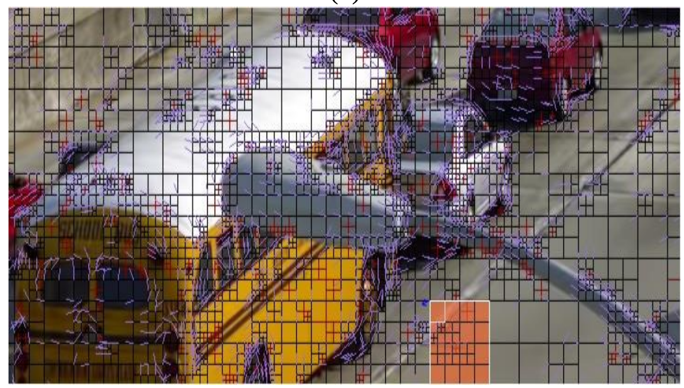
To represent motion vectors and coded block structures for both test sequences and tested configurations Elecard_hevc_analyzer is used, too.

Fig. 6 (a) shows small segment of motion vectors for PeopleonStreet test sequence in lowdelay configuration, while highlighted block represents one example of 64x64 coded block and its structure. Also, Fig. 6 (b) shows small segment of motion vectors for Traffic test sequence in same configuration, as well as one example of 64x64 coded block. Next, Fig. 6 (c) shows small segment of motion vectors for PeopleonStreet test sequence in randomaccess configuration, while highlighted block represents one example of 64x64 coded block. Finally, Fig. 6 (d) shows small segment of motion vectors for Traffic test sequence in same configuration, as well as one example of 64x64 coded block.

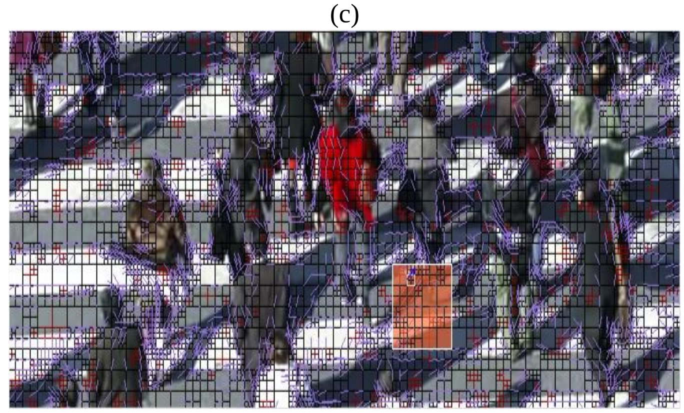
(a)



(b)



(c)



(d)

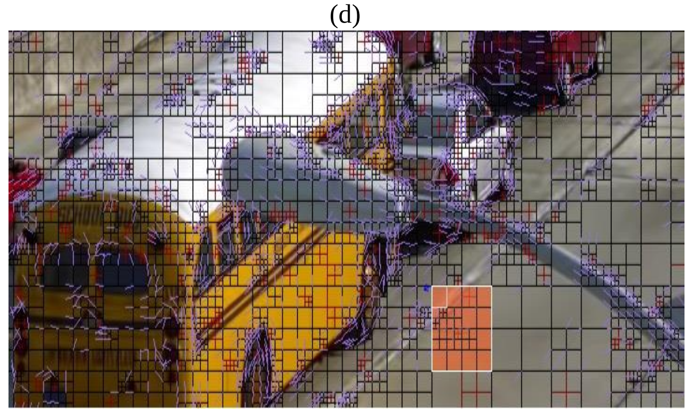


Fig. 6. HEVC motion vectors and coded block stuctures for PeopleOnStreet and Traffic test sequences processed in HM-16.15 reference software.

V. Conclusion

The results presented in this paper indicate HEVC standard HM-16.15 when lowdelay and ranadomaccess configurations are compared. The SNR, bit-rate and encoding time saving are measured for test sequences in same resolution when B picture format (IBBB) is processed. Simulations results have shown that there are small differences in SNR values when both tested sequences are compared i.e. 5% or 6 % depend of configurations. From the bit rate point of view, there are obvious differences in values i.e. the bits rate of test sequence PeopleOnStreet is increased 72% and 74% depending on configuration in comparison with Traffic test sequence. Also, for PeopleonStreet test sequence the encoding time saving is increased 31,50% and 32,86% depending on configuration when that test sequence is compared with Traffic test sequence. Next, results of subjective video assessment for all tested sequences processed by YUV player are provided, when performance for HEVC HM-16.15 encoder are analyses. Finally, percent of intra and inter prediction in total amount encoder compression values is analyzed, as well as motion vectors and coded block structures for both test sequences and tested configurations.

REFERENCES

[1] I.E. Richardson, *Coding Video: A Practical Guide To HEVC And Beyond*, Wiley, 2017.

[2] G. Correa et al. "Performance and computational complexity assessment of high efficiency video encoders", *IEEE Trans. on Circuits and Systems Video Technology*, vol. 11, no.12, pp. 1899-1909, Dec. 2012.

[3] Z. Milicevic, Z. Bojkovic, "HEVC model HM-15,12 vs HM-16.6: objective and subjective performance analysis" *WSEAS Transactions on Signal Processing*, Volume 13, pp. 40-48, 2017.

[4] Z. Milicevic, Z. Bojkovic, K.R.Rao, "HEVC vs H.264/AVC Standard Approach to Coders Performance Evaluation" *WSEAS Transactions on Signal Processing*, Volume 11, pp. 59-63, 2015.

[5] G.J. Sullivan, J. R. Ohm, W. J. Han, and T. Wiegand., "Overview of the High Efficient Video Coding (HEVC) Standard", *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 22, no. 12, pp. 1649-1668, December 2012.

[6] V. Sze, M. Budagavi and G. J. Sullivan (Editors), *High Efficiency Video Coding (HEVC): Algorithms and Architectures,* Springer Cham Heidelberg New York Dordrecht London, Springer International Publishing Switzerland 2014.

[7] svn://hevc.kw.bbc.co.uk/svn/jctvc-hm/, 2017.

[8] J-R. Ohm, G. J. Sullivan, H. Schwarz, T. K. Tan, and T. Wiegand, "Comparison of the Coding Efficiency of Video Coding Standard – Including High Efficient Video Coding (HEVC)", *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 22, no. 12, pp. 1669-1684, December 2012.

Zoran M. Miličević received the B. S. degree in 1996, M. Sc. degree from the University of Belgrade, Republic of Serbia in 2003 and PhD degree from Military Academy Belgrade in 2010 and the University of Belgrade in 2011.
He is the author and co-author of more than 50 scientific journal and conference papers. His current research interest includes video and image coding compression standards, algorithms and implementation.

Jovan G. Mihajlović received the B. S. degree in 2016, and he is working towards M. Sc. degree on Military Academy Belgrade, Republic of Serbia.
He is the author and co-author of 5 conference papers. His current research interest includes video and image coding compression algorithms and implementation.

Zoran S. Bojković with the University of Belgrade, where he is a Professor of Electrical Engineering. He was a visiting professor in USA, Korea, Germany, Norway and many other countries all over the world.
He has published more than 450 invited, regular and tutorial papers in international books, journals and conference proceedings. He is also an active reviewer and a member of scientific committees of numerous journals and conferences. He is co-authors of 7 international books in the area of multimedia communications published by Prentice Hall Wiley and CRC Press. In the last three years he is co-author 4 chapters in international books published by Springer. His areas of interest are image and video processing as well as multimedia communication. He is a life Senior Member IEEE, Member of Serbia Scientific Society, and Engineering Academy of Serbia.