# Performance Comparison of H.265/HEVC, H.264/AVC and VP9 Encoders in Video Dissemination over VANETs

Cristhian Iza Paredes(✉), Ahmad Mohamad Mezher, and Mónica Aguilar Igartua

Universitat Politècnica de Catalunya (UPC), 08034 Barcelona, Spain
cristhian.iza@entel.upc.edu

**Abstract.** Video consumption over VANETs will increase significantly bandwidth and will occupy an important part of the overall data traffic. To decrease the load on the VANET infrastructure and reduce bandwidth taken by video, high efficiency video codecs have been developed. In this work, a benchmarking of H.265/HEVC, H.264/AVC and Google VP9 has been conducted by means of objective and subjective evaluations, assuming an urban VANET scenario. Considering a wide range of bit rates from very low to high, results show a clear advantage of HEVC with average bit rate savings of 27% when compared to VP9 and 49% when compared to AVC.

**Keywords:** HEVC · H.265 · H.264 · AVC · VP9 · VANETs · Video dissemination

## 1 Introduction

Supporting video dissemination over Vehicular Ad Hoc Networks (VANETs) is an attractive feature for many road safety applications. A key component to efficiently transport video with its stringent playout deadlines and bursty traffic characteristics, is using the most-efficient available encoding format. The current video codec standard H.264/AVC provides a better compression efficiency compared to other standards such as H.262/MPEG-2 or VP8. The goal behind the H.264 standard was to provide high quality video at lower bit rates. However, the emerging of a more efficient next generation video coding standard is a high demand at the moment. Two main contenders for the position of the next state of the art video standard are H.265/HEVC [1] and Google VP9 [2]. H.265/HEVC is the latest video coding standard, which achieves an increase of about 50% in coding efficiency compared to its predecessor H.264/AVC [3]. On the other hand, VP9 is an efficient open source video codec developed as part of the WebM Project by Google to get a royalty-free compression standard with efficiency superior to AVC [4].

In our previous work [5,6] H.265/HEVC appeared to provide the best compression efficiency compared to H.264/AVC. In this work we aim to evaluate


O. Gaggi et al. (Eds.): GOODTECHS 2016, LNICST 195, pp. 51–60, 2017.
DOI: 10.1007/978-3-319-61949-1_6

the efficiency of the video compression standards H.265/HEVC and VP9. Our interest is centered on using a video dissemination mechanism in an urban scenario where vehicles' traffic is relatively dense and the communications are more exposed to interferences and radio obstacles.

The rest of the paper is organized as follows: Sect. 2 describes the features of selected encoders. Section 3 discusses the main approach aimed towards an effective solution for video dissemination over VANETs. The performance comparison of encoders and simulation results are discussed and presented in Sect. 4.3. Finally, conclusions and future work are drawn in Sect. 4.4.

## 2 Selected Encoder Implementations

In this section, a brief overview of the selected representative encoders is presented.

**VP9 Encoder.** Google started an Open Source project to develop royalty-free video codecs for the web entitled the WebM Project. The codec developed in the WebM project called is VP9 and is currently being served extensively by Google Chrome and YouTube. To evaluate VP9 compression efficiency, we use the open source libvpx encoder in its version 1.6.0 [2]. It has a two-pass run option which results in the improved rate distortion performance and which is also used in our work.

**H.264/AVC Encoder.** The latest version of JM reference software model (JM 19) was used for encoding video sequences with AVC [7]. The H.264/AVC standard has proven to be very fast, reliable, and efficient. Similarly as VP9, H.264/AVC has a two-step run option. At the first pass, a file with the detailed statistic data about every input frame is generated. At the second step, this information is used to improve the encoder rate-distortion performance.

**H.265/HEVC Encoder.** For evaluating H.265/HEVC-based encoding [1], we selected the latest reference model 16 (HM 16.9) in its simplified model to estimate the compression efficiency of the H.265/HEVC standard. To get constant QP (Quantization Parameter) on each frame we modified *Qpoffset* values of the GOP (Group of Pictures) structure in the configuration file.

The configuration parameters for HEVC, AVC and VP9 were set so that similarity was ensured between the three codecs to avoid any penalization. More details about the configurations can be found in Table 1.

### 2.1 Dataset

The comparison was carried out on the video sequences listed in Table 3. Four video sequences were downloaded from [8] and were used in the simulations, with different spatial, temporal characteristics and frame rates.

**Table 1.** Selected parameters and settings for the AVC, HEVC, and VP9 codecs.

| Codec | Version | Parameters |
|---|---|---|
| HEVC | HM 16.9 | **TAppEncoderStatic** -c encoder_lowdelay_P_main10.cfg (Default main low-delay profile with P frames) -c Traffic.cfg -b encoded sequence.bin -o decoded sequence.yuv -q <QP> |
| AVC | JM 19 | **lencod** -f encoder.cfg -p FrameRate=<FR> -p QPISlice=<QP> -p QPPSlice=<QP> -p QPBSlice=<QP> -p Bitrate=<BR> -p SourceWidth=<W> -p SourceHeight=<H> |
| VP9 | v1.6.0-326 | **vpxenc** --codec=vp9 --profile=0 --fps=<FR> --static-thresh=0 --drop-frame=0 --good --auto-alt-ref=1 --kf-min-dist=8 --kf-max-dist=8 --cq-level=<QP> --max-intra-rate=8 --target-bitrate=<BR> --static-thresh=4 -w <W> -h <H> --limit=500 <inFile>.yuv −o <outFile>.webm |

Each video file was encoded with all three evaluated codecs. Since fixed QP[1] configuration was used to control the quality of AVC, HEVC, and VP9 compressed bitstreams, the sequences were encoded at various QP values trying to cover the full quality scale for each content.

We aim to compare maximum video compression efficiency provided by the latest standards. Based on our previous work [6], we selected Low-Delay-P (LP) coding configuration to reflect the real-time application scenario for all encoders. In this mode the first frame is an intra-frame while the others are encoded as generalized P frames. This makes this mode more vulnerable to packet losses since it needs to wait to receive an entire GoP before decoding the video frames. To mitigate large dependencies between frames and trying to achieve a better packet loss resilience, the GOP size was set to 8 pictures and the Intra Period was set to 25 and 30 pictures for 25 and 30 fps contents, respectively. Table 2 reports the final sets of targeted (R1'–R4') and actual (R1–R4) bit rates, with corresponding QPs, for each codec.

## 3 Video Dissemination in VANET

The realization of a reliable transmission of video over VANETs is extremely challenging mainly due to the network's dynamic topology and stringent requirements of the video streaming service. The high velocity and limited communication range of the vehicles incur frequent link disconnection and even network partition. To evaluate the efficiency of the video compression standards over

[1] The Quantization Parameter (QP) regulates how much spatial detail is saved. When QP is very small, almost all that detail is retained. As QP is increased, some of that detail is aggregated so that the bit rate drops, but at the price of some increase in distortion and some loss of quality.

**Table 2.** Target $R_i$' and actual $R_i$ bit rates (kbps) including the corresponding QP values for each codec.

| Sequence | Codec | R1' | R1 | QP | R2' | R2 | QP | R3' | R3 | QP | R4' | R4 | QP |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Highway | AVC | 375 | 384 | 30 | 750 | 747 | 24 | 1500 | 1574 | 15 | 2500 | 2515 | 11 |
| | HEVC | 375 | 336 | 27 | 750 | 776 | 24 | 1500 | 1450 | 21 | 2500 | 2717 | 18 |
| | VP9 | 375 | 390 | 28 | 750 | 749 | 25 | 1500 | 1486 | 22 | 2500 | 2833 | 19 |
| Hall monitor | AVC | 375 | 385 | 32 | 750 | 779 | 25 | 1500 | 1590 | 17 | 2500 | 2877 | 13 |
| | HEVC | 375 | 363 | 28 | 750 | 675 | 25 | 1500 | 1319 | 22 | 2500 | 2452 | 19 |
| | VP9 | 375 | 416 | 30 | 750 | 787 | 26 | 1500 | 1640 | 22 | 2500 | 2370 | 20 |
| City | AVC | 256 | 242 | 58 | 512 | 520 | 33 | 1024 | 1010 | 23 | 2048 | 2087 | 12 |
| | HEVC | 256 | 235 | 35 | 512 | 508 | 29 | 1024 | 1392 | 24 | 2048 | 2041 | 19 |
| | VP9 | 256 | 253 | 37 | 512 | 535 | 31 | 1024 | 1126 | 25 | 2048 | 2195 | 20 |
| Bus | AVC | 256 | 251 | 54 | 512 | 539 | 43 | 1024 | 1006 | 34 | 2048 | 2038 | 22 |
| | HEVC | 256 | 248 | 40 | 512 | 514 | 34 | 1024 | 997 | 29 | 2048 | 2089 | 23 |
| | VP9 | 256 | 267 | 41 | 512 | 512 | 36 | 1024 | 1080 | 30 | 2048 | 2192 | 24 |

VANETs, we use a smart dissemination protocol known as RCP+ that we proposed in previous works [5,6]. The proposed mechanism is built on top of IEEE 1609.3 by adding a layer to select next forwarder vehicles based on the information of the environment and an estimation of the congestion of the communication channel. RCP+ ensures a large dissemination in the network to rebroadcast the video content.

### 3.1 Scenario Description

We focus the situation on the immediate consequences of a traffic accident. The crashed vehicle starts to generate and transmit a real-time SOS message to alert the vehicles in the network about the incident and to the appropriate emergency centers (e.g. 112 or 911). The emergency message includes a short video of a few seconds before the crash. We consider a real street environment imported from OpenStreetMap [9]. Under the street model, vehicles are generated and their moving patterns are controlled by SUMO [10]. Shadowing models are used to reproduce the attenuation of a radio signal induced by obstacles, such as

**Table 3.** Test video sequences have a resolution of $352 \times 288$ pixels

| Sequence | Frame rate | Number of frames |
|---|---|---|
| Highway | 25 fps | 2000 |
| Hall monitor | 30 fps | 300 |
| City | 30 fps | 300 |
| Bus | 25 fps | 150 |

**Table 4.** Simulation parameters.

| | Parameter | Value |
|---|---|---|
| Physic and MAC Layers IEEE 802.11p | Channel; Bandwidth | 178, 5.89 GHZ; 10 MHz |
| | Transmission range | 230 m |
| | Transmission power | 20 mW |
| | Obstacle model | Defined in [11,12] |
| | Beacon [$CW_{min}$, $CW_{max}$], AIFSN | [15,1023], 6 |
| | Data [$CW_{min}$, $CW_{max}$], AIFSN | [7,15], 3 |
| | Bit rate | 6 Mbit/s |
| RCP+ [5,6] | $RSS_{th}$, $RSS_{max}$ | −89 dBm, −20 dBm |
| | Time slot | 13 µs |
| | Time window | 10 s |
| | $\delta$ (Waiting Time) | [1, 11] µs |
| | Beacon frecuency, Beacon size | 1 Hz, >=32 bytes |
| Scenarios | Number of runs per point | 10 |
| | Time to live (TTL) | 90 s |

buildings or other structures blocking the direct line of sight. A set of 4 RSUs (Road Side Units) have been strategically located at 20 m, 300 m, 600 m, and 1200 m from the accident scene. The distance between the RSUs and the road is 3 m. RSUs are traffic sinks used to measure the quality of the received video at different distances from the accident.

## 4 Performance Evaluation

This section provides simulation results on the coding performance of the three video coding standards under evaluation. We first present the simulation setup used, including models and scenarios. Then, we present the comparison of the compression efficiency between HEVC, VP9 and AVC by means of objective and subjective evaluations in the considered VANET video streaming scenario.

### 4.1 Simulation Setup

To carry out the performance comparison, each run uses a different random scenario that fulfills the requirements of the study. For each point in all figures we have calculated the average from 10 simulation runs. This let us obtain a standard error less than 5% in a 95% confidence interval. The packet error and Medium Access Control (MAC) layer models adopted are based on the IEEE 802.11p, using a data rate of 6 Mbit/s, a transmission power of 20 mW, and a receiver sensitivity of −89 dBm. In addition, all hello messages use the same Access Category (AC_BE), thus with the same values of Contention Window

(CW) and Arbitration Inter-Frame Spacing (AIFSN). Table 4 contains a summary of the simulation parameters common to all simulation scenarios.

We assume that each vehicle is equipped with a GPS device to obtain its geographical location in current time. A preloaded digital map provides information about roads. We assume that vehicles periodically exchange their own physical location, moving velocity and direction information enclosed in their periodic hello messages. They are sent at the frequency of 1 Hz. Finally, vehicles are assumed to be equipped with IEEE 802.11p wireless technology and computation capabilities.

### 4.2 Performance Measures

We use three performance metrics to evaluate the quality of video transmitted over VANETs:

*Frame Delivery Ratio*: It is defined as the ratio between the number of frames delivered and the total number of frames received during a time interval of $T$ seconds.

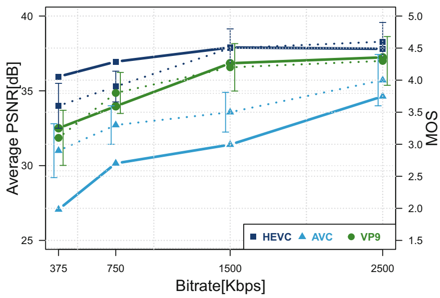
*PSNR*(*Peak Signal-to-Noise Ratio*): It is an objective metric used to assess the application-level QoS of video transmissions. PSNR measures the error between the reconstructed image and the original one, frame by frame. We assume that in case an individual frame was lost, the decoder would display the last successfully received frame of the same type. So if a frame is dropped, we need to compare the source frame to the previous streamed frame.

*MOS*(*Mean Opinion Score*): It is a subjective metric used to provide a numerical indication of the perceived quality from the users's point of view of the received video. In a MOS assessment test, video sequences are presented in a predefined order to a group of subjects, who are asked to rate their visual quality on a rating scale. The MOS score is expressed in the range from 1 to 5, where 5 is the highest perceived quality and 1 is the lowest perceived quality.

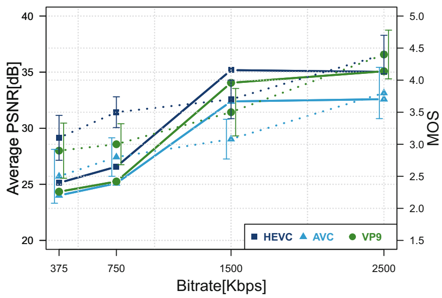
### 4.3 Results and Discussion

In a first set of experiments, we used the Bj$\phi$ntegaard model [13] to calculate the coding efficiency between different codecs. This metric allows us to compute the average gain in PSNR or the average per cent saving in bitrate between two rate-distortion curves. Also, we used another model based on subjective quality scores [14]. This model computes the average MOS difference and average bit rate difference between two sets of subjective results corresponding to two different codecs. This model reports the average bit rate difference, $\Delta$R, for a similar perceived visual quality. Table 5 provides the results in terms of BD-Rate[2] and $\Delta$R results. Results based on the Bj$\phi$ntegaard model show that the average bit rate reduction of HEVC relative to AVC and VP9 is 49.73% and 27.12%,

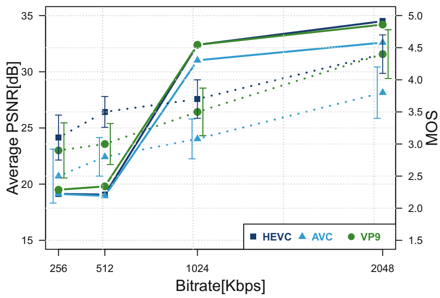
[2] Bj$\phi$ntegaard Delta-Rate (BD-Rate) is the average bit rate difference in percentage for the same PSNR.



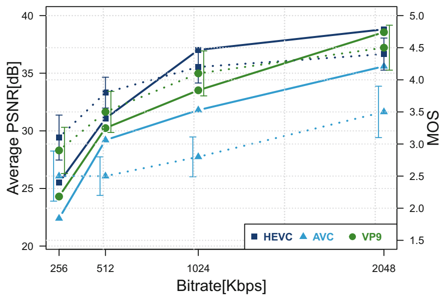

(a) Highway, 348x288, 25 fps




(b) Hall Monitor, 348x288, 30 fps




(c) City, 348x288, 30 fps



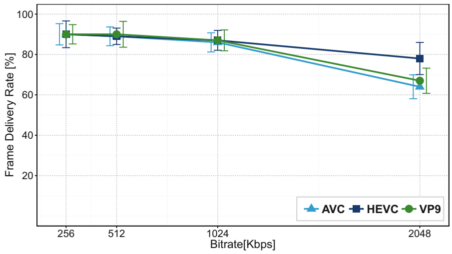

(d) Bus, 348x288, 25 fps

**Fig. 1.** PSNR (solid line) curves and subjective MOS (dashed line) values, for each bit rate and each video content. 95% confidence intervals are shown.

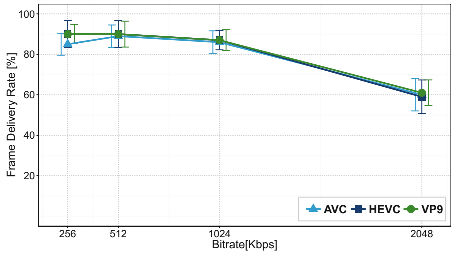
**Table 5.** Comparison of the three evaluated coding algorithms in terms of bit rate reduction for similar PSNR and MOS. Negative values indicate actual bit rate reduction.

| Sequence | HEVC vs AVC | | VP9 vs AVC | | HEVC vs VP9 | |
|---|---|---|---|---|---|---|
| | BD-Rate | $\Delta$R | BD-Rate | $\Delta$R | BD-Rate | $\Delta$R |
| Highway | −47.41% | −40.11% | −32.48% | −36.58% | −41.19% | −42.79% |
| Hall monitor | −32.60% | −23.70% | −27.38% | −20.08% | −9.57% | −12.80% |
| City | −51.11% | −47.01% | −42.89% | −34.29% | −21.66% | −26.16% |
| Bus | −67.82% | −65.62% | −52.64% | −50.44% | −36.05% | −39.25% |
| Average | −49.73% | −44.11% | −38.85% | −35.35% | −27.12% | −30.25% |

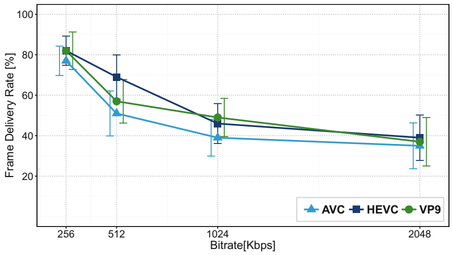
respectively. Also, the average bit rate reduction of VP9 relative to AVC is 38.85%. On the other hand, results based on the subjective ratings indicate an average bit rate saving of 44.11% and 30.35% for HEVC when compared to AVC and VP9, respectively. Furthermore, the bit rate reduction achieved by VP9 relative to AVC is 35.35%. As it can be seen, HEVC encoder provides better results than all the other codecs avaluated.



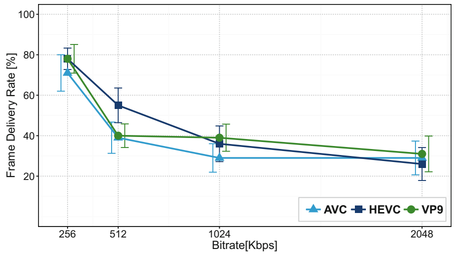

(a) $RSU_1$ located 20 m. from accident




(b) $RSU_2$ located 300 m. from accident

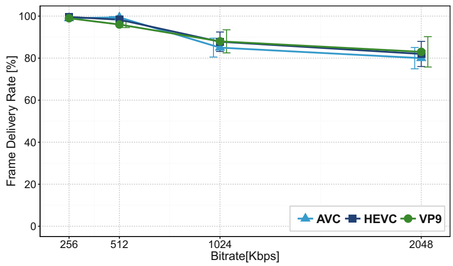



(c) $RSU_3$ located 600 m. from accident

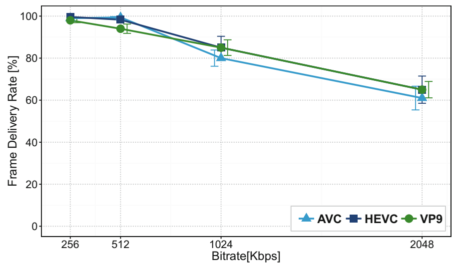



(d) $RSU_4$ located 1200 m. from accident

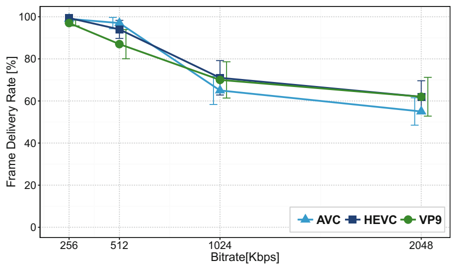
**Fig. 2.** Urban medium density scenario: 60 vehicles/km$^2$. Frame delivery rates with 95% confidence intervals for the CITY.



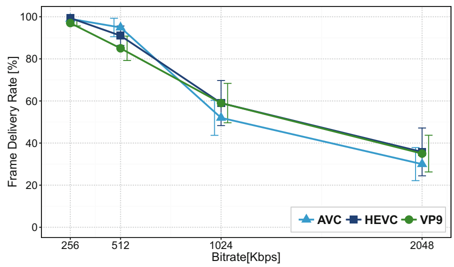

(a) $RSU_1$ located 20 m. from accident




(b) $RSU_2$ located 300 m. from accident




(c) $RSU_3$ located 600 m. from accident




(d) $RSU_4$ located 1200 m. from accident

**Fig. 3.** Urban high density scenario: 120 vehicles/km$^2$. Frame delivery rates with 95% confidence intervals for the CITY.

As a next step, we carry out a comparative assessment for the Low-Delay-P (LP) configuration of H.265/HEVC, VP9, and H.264/AVC encoders. Figure 1 shows the Rate-Distortion curves based on PSNR measurements and subjective ratings based on MOS measurements for all sequences. Based on PSNR measurements, HEVC outperforms VP9 by 0.5 to 3.5 dB, while VP9 provides a gain ranging from 0.5 to 8.45 dB when compared to AVC. For all video contents and bit rates, objective measurements show that HEVC outperforms both VP9 and AVC coding algorithms. The subjective results show similar trend to objective measurements: HEVC provides the best visual quality for a similar bit rate and outperforms AVC in most cases. Also, VP9 achieves better visual quality than AVC. However, in some cases (in particular, at high bit rates), HEVC and VP9 have similar ratings and there is no sufficient statistical evidence indicating differences in performance between these codecs at these bit rates.

Finally, we compare the effectiveness of the RCP+ scheme in terms of frame delivery rate for each codec. In the urban scenario, we define three densities: 30, 60 and 120 vehicles/km$^2$. These densities can be considered as Low, Medium, and High densities of vehicles. These network densities cover a range from low (normal or night time) to high vehicular traffic density (rush hour). A vehicle operating in a sparse traffic density is said to be in a totally disconnected neighborhood if it has no vehicle neighbor within its transmission range. In this case, simulation results (not shown here due to space limits) indicate that only the $RSU_1$ and $RSU_2$ located 20 and 300 m from the accident, received the complete trace. This makes it difficult to evaluate the codec in this scenario. On the other hand, the performance of our mechanism in Medium and High densities are presented in Figs. 2 and 3, respectively. As it is clearly seen, the HEVC encoder provides gains in terms of Frame Delivery Ratio compared to both VP9 and AVC encoders. Also, as the distance from the accident increases for the RSU, the delivery ratio decreases since probability of collisions or network failure increases. This result is expected, because the urban scenario shows more aggressiveness in the packet loss due to the existence of buildings. Besides, dynamic topology networks in VANET causes temporary disconnections, interrupting the video message dissemination and compromising the delivery of the video frames.

### 4.4 Conclusion

In this paper we have studied compression efficiency of the current video compression standard and candidates for the next generation video coding standard over VANETs in an urban traffic scenario. The high bandwidth required for video dissemination can be tackled through the use of recent encoders that allow doubling the efficiency coding, reducing almost half the bit rate for similar levels quality. The results have shown the superior compression efficiency of H.265/HEVC coding standard over H.264/AVC and VP9 encoders. The possible drawback of using H.265/HEVC is a higher computational complexity. As future work, we will seek an efficient forwarding mechanism to enhance video QoE (Quality of Experience) according with the VANET safety applications' requirements.

**Acknowledgments.** This work was partly supported by the Spanish Government through projects TEC2014-54435-C4-1-R (INcident monitoRing In Smart COmmunities. QoS and Privacy, INRISCO) and AGAUR Information Security Group (ISG) project - 2014 SGR 1504. Cristian Iza Paredes is recipient of a grant from Secretaria Nacional de Educación Superior, Ciencia y Tecnología SENESCYT. Ahmad Mohamad Mezher is a postdoctoral researcher with the Information Security Group in the Department of Network Engineering at the Universitat Politècnica de Catalunya (UPC). Additional funding supporting this work has been granted to UPC by the Spanish Ministry of Economy and Competitiveness (MINECO) through the "Anonymized Demographic Surveys (ADS)" project, ref. TIN2014-58259-JIN, under the funding program "Proyectos de I+D+i para Jóvenes Investigadores".

## References

1. High efficiency video coding (HEVC). https://hevc.hhi.fraunhofer.de/. Accessed 17 May 2016
2. The webm project. http://www.webmproject.org/. Accessed 5 Sep 2016
3. Grois, D., Marpe, D., Mulayoff, A., Itzhaky, B., Hadar, O.: Performance comparison of h.265/MPEG-HEVC, vp. 9, and H.264/MPEG-AVC encoders. In: Proceedings of the Picture Coding Symposium (PCS), pp. 394–397, December 2013
4. Sharabayko, M.P., Markov, N.G.: Contemporary video compression standards: H.265/HEVC, vp9, vp10, daala. In: Proceedings of the International Siberian Conference Control and Communications (SIBCON), pp. 1–4, May 2016
5. Iza-Paredes, C., Mezher, A.M., Aguilar Igartua, M.: Evaluating video dissemination in realistic urban vehicular ad-hoc networks. In: The 19th ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems, MSWiM 2016, Malta. ACM (2016)
6. Iza-Paredes, C., Mezher, A.M., Aguilar Igartua, M.: Adaptive video-streaming dissemination in realistic highway vehicular ad-hoc networks. In: Performance Evaluation of Wireless Ad Hoc, Sensor, and Ubiquitous Networks, PE-WASUN 2016, Malta. ACM (2016)
7. H.264/avc software coordination (AVC). http://iphome.hhi.de/suehring/tml/. Accessed 17 May 2016
8. Derf's dataset. https://media.xiph.org/video/derf/. Accessed 22 Mar 2016
9. OpenStreetMap. http://www.openstreetmap.org. Accessed 22 Mar 2015
10. SUMO – Simulation of Urban MObility. http://goo.gl/uvvD4N. Accessed 22 Mar 2015
11. Boban, M., Vinhoza, T.T.V., Ferreira, M., Barros, J., Tonguz, O.K.: Impact of vehicles as obstacles in vehicular ad hoc networks. IEEE J. Sel. Areas Commun. **29**(1), 15–28 (2011)
12. Sommer, C., Joerer, S., Segata, M., Tonguz, O.K., Cigno, R.L., Dressler, F.: How shadowing hurts vehicular communications and how dynamic beaconing can help. IEEE Trans. Mob. Comput. **14**(7), 1411–1421 (2015)
13. Bj$\phi$ntegaard, G.: Calculation of average PSNR differences between RD-curves. ITU-T SG16/Q6, Austin, Texas, USA (2001)
14. Hanhart, P., Ebrahimi, T.: Calculation of average coding efficiency based on subjective quality scores. J. Vis. Commun. Image Represent. **25**(3), 555–564 (2014). QoE in 2D/3D Video Systems