Video transmission over IEEE 802.11p: real-world measurements

Alexey Vinel, Senior Member, IEEE, Evgeny Belyaev, Member, IEEE, Olivier Lamotte, Moncef Gabbouj, Fellow, IEEE, Yevgeni Koucheryavy, Senior Member, IEEE and Karen Egiazarian, Senior Member, IEEE

Abstract—IEEE 802.11p/ITG-G5 vehicle-to-vehicle communication technology, which enables the new class of safety and infotainment applications, is currently an emerging research topic in both industry and academia. The proposed spectrum allocation of 10 Mhz channels for DSRC (Dedicated Short Range Communication) in 5.9 GHz band for the USA and Europe allows considering the transmission of video information between vehicles as one of the grounding blocks for future automotive applications. Although several published works addressed the problems of video content delivery in VANETs (Vehicular Ad-hoc NETworks), no work has been reported on real-world measurements of visual quality for video being transmitted over the IEEE 802.11p vehicle-to-vehicle communication channel. This paper presents a real-time scalable video codec as well as the first results of visual quality measurements for the video information transmitted using the off-the-shelf Componentality FlexRoad DSRC equipment.

Index Terms—Video applications, vehicle-to-vehicle communications, test-bed, real-world measurements, VANET, IEEE 802.11p.

I. INTRODUCTION

The coming years will see the adoption of IEEE 802.11p/ITG-G5 technology, which enables broadband vehicle-to-vehicle and vehicle-to-roadside connectivity. The design and validation of this technology as well as prospective safety and infotainment applications in VANETs (Vehicular Ad-hoc NETworks) are currently areas of intensive research. The availability of 10 Mhz channels for DSRC (Dedicated Short Range Communication) in 5.9 GHz band allows considering video information transmission between vehicles as a feasible opportunity for new use cases of the 802.11p technology.

A lot of recent research efforts has been dedicated to the efficient video-streaming in VANETs. A novel application-centric multi-hop routing protocol for video transmission is proposed in [2] to maximize the end users’ peak signal-to-noise ratio. Multi-source streaming across a VANET overlay network is studied in [3], while video streaming from one source to all nodes in an urban scenario is analyzed in [4].

Dynamic service schemes, which maximize the total user-satisfaction and achieve an appropriate amount of fairness in peer-to-peer-based VANETs, are proposed in [5]. Efficient channel adaptation mechanism for real-time vehicle-to-vehicle video transmission is analyzed in [6].

The commonality of the above reviewed works is that they all present either analytic or simulation based work. In this paper we report real-world experiments with FlexRoad DSRC video test-bed. Although the considered scenario is limited to the case of one communicating vehicle-pair, it will attract more attention to this topic and speed up the practical assessment studies of new generation video-based vehicular systems.

The contributions of the paper can be summarized as follows:

• the developed video test-bed, which includes FlexRoad DSRC platform and scalable video codec based on three-dimensional discrete wavelet transform (3-D DWT), is described;
• the approach to measure expected end-to-end visual quality of transmitted video information is proposed;
• some preliminary experimental results are reported.

The paper is organized as follows. General structure of the considered system as well as the examples of prospective VANET-enabled video applications are provided in Section II. The description of the developed test-bed components is given in Section III. Section IV reports the experimental results, while Section V concludes the paper.

II. VIDEO APPLICATIONS IN VANETS

The considered video-based vehicular test-bed includes the following main components (see Figure 1):

• IEEE 802.11p transceiver;
• video compressing device;
• video cameras.

With this set of equipment and depending on the area, where the cameras are directed to, i.e. inside or outside of the vehicle, numerous applications can be implemented for the goals of public security, road safety and traffic control. Here we give three examples of such applications:

• overtaking assistance;
• in-vehicle video surveillance;
• traffic conditions video surveillance.
Let us describe these proposed applications in more details.

A. Overtaking assistance

In an overtaking assistance application a video stream captured by a windshield-mounted camera in a vehicle is compressed, broadcast to the vehicle driving behind it and displayed to its driver. Such a "see-through" system is aimed at helping drivers overtake long and vision-obstructing vehicles, such as trucks, on rural roads using the oncoming lane. Moreover, dangerous road situations or even rear-end collisions can be avoided when information about the obstacle is provided to the driver well in advance, following the observation of the vision vehicle [6].

B. In-vehicle video surveillance

In-vehicle video surveillance application captures a video data by an internal cabin-mounted camera in a vehicle. After compression this information is transmitted to the security services like police, ambulance, etc. Such an application will allow real-time monitoring of public transportation aiming at preventing terrorism, vandalism and other crimes. The efficiency of an in-vehicle video surveillance can be improved by means of video data analysis and detection of the malicious activity using state-of-the-art video analytics methods [7].

C. Traffic conditions video surveillance

For the goal of traffic control that might be needed to see the current situation at a given road segment, particular intersection or even a given traffic lane. Taking the benefits of global positioning system, traffic management center can activate the external cameras of vehicles located in the geographical area of interest. Video information with the current road views is then compressed at vehicle side and transmitted back to the management center. Real-time reaction to traffic jams caused by the accidents can be achieved if traffic conditions video surveillance system is combined with the eCall [8] or a similar system, which automatically informs the emergency services about the crash.

III. TEST-BED COMPONENTS

A. IEEE 802.11p platform

We used Componetality FlexRoad equipment in our test-bed. Componetality Oy, Finnish corporate designing and manufacturing automotive multimedia and communication systems, produces VANET modules for vehicle to vehicle and vehicle to infrastructure communication.

Componetality FlexRoad is a hardware and software platform widely based on open solutions and accepting usage of the external applications running right on the CPU embedded to the connectivity device. FlexRoad also uses Linux operating system, automotive mesh technology support and dedicated tools to provide reliable multifunctional data network over a set of moving vehicles and infrastructural objects.

The characteristics of Componetality FlexRoad devices, used in the paper, are the following:

- 300 MHz Atheros CPU with MIPS architecture;
- 64 MB of RAM;
- 5.9 GHz IEEE 802.11p radio module1;
- 12 dBm and 4 dBm antennas on transmitter and reciever sides respectively.

1This information was provided to us by Componetality Oy. We have not performed a study about the compliance of FlexRoad DSRC devices with the IEEE 802.11p standard.
B. Video codec

In this paper a real-time scalable video codec proposed in [9] is used for video compression. The main idea of the scalable video coding (SVC) is that the encoder forms the bit stream from several layers: the base layer and the enhancement layers. Each layer is characterized by its own bit rate and visual quality. There exist the following types of scalability:

- Temporal scalability (different frame rates);
- Spatial scalability (different frame resolutions);
- Visual quality scalability.

In comparison to traditional coding the scalable coding has several advantages.

- Each receiver can decode only part of the video stream with less frame resolution, frame rate or visual quality depending on computation complexity restrictions or a situation in a channel. For example, in case of video transmission from one vehicle to another, such type of coding allows simultaneously storing of a full resolution video on a vehicle storage device and transmitting a low resolution video, if the channel capacity is low.

- In combination with unequal loss protection (ULP) of different video stream layers [10], SVC provides robust transmission even for channels with high packet losses. For example, ULP can be achieved based on inter-packet Reed-Solomon (RS) codes for channel with erasures and implemented on application network layer. The base video stream layer is protected using RS codes with high a redundancy level while the remaining layers are protected with a lower redundancy level or not protected at all. If the delivery of the base layer is provided, then even for high packet loss probabilities a basic visual quality on the receiver side will be guaranteed.

A general scheme of the scalable video codec is shown in Figure 2. This codec is based on 3-D DWT and includes the following main coding stages. First, a group of frames (GOF) of length 16 are accumulated in the input frame buffer. Then wavelet transformation is applied: one-dimension Haar wavelet transform of length 16 in the temporal direction and 5/3 spatial lifting wavelet transform at three-levels of the decomposition for each frame in the GOF. Then, each wavelet subband is independently compressed using bit-plane entropy coding. For the required bit rate achievement a rate controller selects the quantization step (or truncation point) for each subband. For more details see [9].

In case of independent wavelet subband coding, each subband can be decoded even if one or more subbands are not delivered. If some subband is not received, then the reconstructed video will have a lower quality: some horizontal, vertical or diagonal details may be lost. With the increase in packet loss probability, more and more subbands will not be correctly received, but the video will be still recognizable until the main low-frequency subband (which corresponds to the base layer) in the GOF is delivered (see the example in Figure 5d). So, such type of coding scheme with high protection of low-frequency subbands can provide acceptable video quality even for high packet loss probabilities.

C. Expected end-to-end visual quality measurement

For a expected end-to-end visual quality estimation we use the following approach (see Figure 3). During each measurement, a test video sequence is compressed by the 3-D DWT encoder. Then each wavelet subband bit stream is placed in one RTP packet and transmitted over DSRC channel (without any ULP). At the receiver side the RTP packets are accumulated in a buffer and, then, the 3-D DWT decoder starts decompressing and playback. The reconstructed video is stored at the receiver hard disk. Then, for measurement with number i the sum of square errors is calculated:

\[
D_i = \frac{1}{F \cdot W \cdot H} \sum_{f=0}^{F-1} \sum_{x=0}^{W-1} \sum_{y=0}^{H-1} (s[x, y, f] - s'[x, y, f])^2,
\]

where \(s[x, y, f]\) and \(s'[x, y, f]\) are the luma values for pixel with coordinates \((x, y)\) in frame \(f\) for the original and reconstructed video sequences, respectively, \(W\) is the frame with, \(H\) is the frame height, \(F\) is the number of frames in the test video sequence.

Then we repeat the experiment \(N\) times and estimate the Expected Peak Signal-to-Noise Ratio (PSNR) which we use as the visual quality metric:

\[
E[PSNR] = 10 \log \frac{s_{max}^2}{\frac{1}{N} \sum_{i=0}^{N-1} D_i},
\]

where \(s_{max} = 255\) is the maximum possible luma value.

IV. EXPERIMENTAL RESULTS

The experiments were handled in a major street close to the university campus in Hervanta, a suburb of Tampere, Finland. At the time of the experiments, the traffic was moderate and some cars and buses occasionally obstructed the line of sight between the transmitter and the receiver vehicles.

In our work, a test video sequence contains the captured video from the moving vehicle on the road (see Figure 5a) with parameters \(W = 640\), \(H = 480\) and \(F = 300\). The capturing frame rate is 30 frames per second. The expected visual quality measurements for the inter-vehicle distances \(d = \{2, 100, 200, 300, 400, 500, 600, 700, 800\}\) meters were performed. For each distance \(N = 9\) measurements described in Section III-C were done. The required bit rate was set to 1200 kbps, what provides a DVD-like visual quality (around 35dB) when all video packets are delivered.

Figure 4 shows the estimated expected visual quality depending on the inter-vehicle distance. Figure 5 shows some examples of video frames. One the one side, the proposed test-bed provide a good visual quality if the inter-vehicle distance is less than 600–700 meters. If the the inter-vehicle distance is more than 700 meters, then packet loss rate becomes unacceptably high for video data delivery. On the other side,
Fig. 2. Video compression based on three-dimensional discrete wavelet transform

Fig. 3. The proposed expected end-to-end visual quality estimation

Fig. 4. Expected end-to-end visual quality depending on distance

V. CONCLUSION

Video transmission over IEEE 802.11p VANETs makes it possible to introduce a variety of new automotive applications, such as the overtaking assistance, in-vehicle video surveillance, etc. Initial practical evaluation results are quite encouraging and allow concluding that good visual quality can be achieved with commercially available Componentality FlexRoad DSRC devices and our 3-D DWT video codec when the inter-vehicle distance is less than 300–400 meters. This is sufficient for applications requiring efficient and real-time short range video delivery.

Future work will be dedicated to the enhancing of video codec by implementing additional measures to increase the reliability of vehicle-to-vehicle video transmission by introducing a real-time unequal packet loss protection based on Reed-Solomon codes and packets interleaving.

ACKNOWLEDGEMENTS

The authors thank Componentality Oy (http://www.componentality.com/) for the provided DSRC FlexRoad platform.

This work has been done with the support of the Finnish Funding Agency for Technology and Innovation (TEKES) under the TIVIT TEKES SRA IoT – Strategic Research Agenda on Internet of Things project.

our experiments show that the continuous video playback is provided only when the inter-vehicle distance is less than 300–400 meters, otherwise the proposed test-bed cannot provide continuous playback due to burst packet losses caused by obstruction of the line of sight between the transmitter and the receiver by other vehicles.
Fig. 5. Reconstructed frame depending on inter-vehicle distance $d$

REFERENCES


