

An Overtaking Assistance System Based on Joint Beacons and Real-Time Video Transmission

Alexey Vinel, *Member, IEEE*, Evgeny Belyaev, Karen Egiazarian, *Senior Member, IEEE*, and Yevgeni Koucheryavy, *Senior Member, IEEE*

Abstract—Overtaking on rural roads often becomes dangerous when oncoming traffic is detected by the driver too late or its speed is underestimated. Recently proposed cooperative overtaking assistance systems, which are based on Vehicular Ad hoc NETWORKS (VANETs), rely on either real-time video transmission or the exchange of status messages (beacons). In the first case, a video stream captured by a camera installed at the windshield of a vehicle is compressed and broadcast to any vehicles driving behind it, where it is displayed to the driver. In the second case, beacons that include position, speed, and direction are frequently broadcast by all the vehicles to ensure detection of oncoming traffic as early as possible and to issue a warning to the driver whenever needed. In this paper, we demonstrate that the performance of a video-based overtaking assistant can be significantly improved if codec channel adaptation is undertaken by exploiting information from the beacons about any forthcoming increase in the load of the multiple access channel used. The theoretical framework presented describes the basic patterns of such a coupled overtaking assistant and can serve as a useful guideline for the future practical implementation of the system. The benefits of our approach are demonstrated in relation to the practical scenario of H.264/AVC video coding and IEEE 802.11p/Wireless Access in Vehicular Environments (WAVE) intervehicle communication standards.

Index Terms—Cooperative awareness, IEEE 802.11p/Wireless Access in Vehicular Environments (WAVE), overtaking assistance, real-time video transmission, Vehicular Ad hoc NETWORK (VANET).

I. INTRODUCTION

OVERTAKING of long and vision-obstructing vehicles such as trucks, using the oncoming lane, is a challenging task. Thus, any information that increases drivers' awareness about the road situation and supports their decision on whether to start the overtaking maneuver is very useful. Often, truck drivers help the drivers of vehicles traveling behind them by actuation of turn indicators. A concept of Vehicular Ad hoc NETWORKS (VANETs) allows extension of such *manual* assistance and enables *automated* cooperation between different vehicles on the road.

Manuscript received December 23, 2011; revised February 18, 2012; accepted March 21, 2012. Date of publication March 27, 2012; date of current version June 12, 2012. The review of this paper was coordinated by Prof. A. Chatterjee.

A. Vinel and Y. Koucheryavy are with the Department of Communication Engineering, Tampere University of Technology, 33101 Tampere, Finland (e-mail: alexey.vinel@tut.fi, yk@cs.tut.fi).

E. Belyaev and K. Egiazarian are with the Institute of Signal Processing, Tampere University of Technology, 33101 Tampere, Finland (e-mail: evgeny.belyaev@tut.fi, karen.egiazarian@tut.fi).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TVT.2012.2192301

The IEEE 802.11p/Wireless Access in Vehicular Environments (WAVE) standard has recently issued a set of physical (PHY) and medium access control (MAC) layer specifications to enable communications in VANETs [1], [2]. In WAVE, each vehicle periodically switches on a common control channel (CCH) to receive and transmit control and safety-related messages and then tune into one of available service channels (SCHs) to exchange all the other information. In both the CCH and the SCH, IEEE 802.11p Enhanced Distributed Channel Access (EDCA), which is a multiple-access technique, is applied, which exploits carrier sensing, as well as using an exponential backoff collision resolution algorithm.

Novel vehicular safety applications mostly rely on beaconing. *Beacons* are the messages carrying status information about the vehicle position, its speed, and direction, and they are periodically broadcast on the CCH and are used to announce the presence of a vehicle to its neighbors. *Event-driven* messages can also be transmitted on the CCH to warn the vehicles in the vicinity about some happening.

The *beacon-based* overtaking assistance system is studied by simulation in [3], where it is shown that, by setting up an adequate beaconing frequency and properly choosing the IEEE 802.11p channel access priorities, the oncoming traffic can be detected early enough to issue a timely warning to the driver.

Information from the overtaking assistant system can be even more relevant if it is presented as *real-time video data*. Much research effort is dedicated to the provision of efficient video streaming in VANETs. A novel application-centric multihop routing protocol for video transmission is proposed in [4] to maximize the end users' peak signal-to-noise ratio. Multisource streaming across a VANET overlay network is studied in [5], whereas video streaming from one source to all nodes in an urban scenario is analyzed in [6]. 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 [7].

In the aforementioned works, it is assumed that a large buffering latency of 5–15 s is acceptable. In contrast, in this paper, we consider the case where a video of the vehicle to be overtaken is captured, encoded, transmitted across the wireless channel, decoded, and displayed in the overtaking vehicle (see Fig. 1). Similar to *video conference* applications, the end-to-end latency between video capturing and displaying should not exceed 200 ms during the overall system operation; otherwise, displayed information becomes outdated. Moreover, acceptable video quality has to be guaranteed.

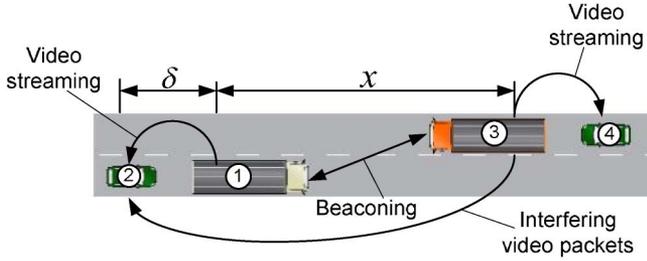


Fig. 1. Overtaking scenario studied.

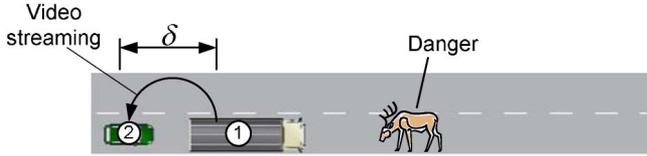


Fig. 2. Sudden danger scenario studied.

In [8], the implementation of video transmissions between the overtaken and overtaking vehicles is introduced. This allows the driver of the overtaking vehicle to have the same view of the road as that of the driver of the overtaken vehicle. However, [8] studies the “ideal” scenario when there are only two vehicles in the network, which is only valid in the early stages of WAVE penetration.

If some designated SCH is exploited for real-time video transmission, then, during a *short interval* of time when two platoons of vehicles are about to meet each other, the SCH load *rapidly increases*. Thus, the main challenge for a video-based overtaking assistance system with a high level of VANET-enabled vehicle penetration is to provide reliable operation in terms of end-to-end latency and the visual quality during the approach of the oncoming platoon. This is in contrast with traditional real-time video transmission, where one is interested in the *average* video quality over a *long* time period.

In practice, to provide low-latency video transmission, one has to *estimate* the channel throughput in real time and to use this information at the encoder side, adapting the video bit rate to varying channel conditions. In cases where the channel estimation algorithm operates with no knowledge about future network conditions, it is difficult to transmit a video without any visual quality degradation, at least over a short time interval since any estimation itself requires some time.

In this paper, we propose a joint overtaking assistance system, which relies on the fact that the information about the oncoming traffic obtained from the beacons and received on the CCH is available on the encoder *before* the SCH degradation occurs. This way, the system is able to provide low end-to-end latency and acceptable visual quality, even during the period when platoons are approaching, which is the most critical time from the overtaking assistant application perspective.

The presented framework is also applicable to the efficient implementation of other novel automotive safety applications, which are based on real-time video transmission. An example of such a use is shown in Fig. 2. As one can see, 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 observation of the vision vehicle.

The *contributions* of this paper are summarized as follows.

- 1) We introduce the novel concept of a coupled overtaking assistance system based on beaconing and real-time video transmission.
- 2) We develop a mathematical framework for evaluation of the performance of the system.
- 3) We characterize the benefits of the proposed approach for the H.264/AVC video coding and IEEE 802.11p/WAVE communication standards.

This paper is organized as follows: In Section II, we describe the system model, including the principal assumptions about the intervehicle communication and the end-to-end video distortion. In Section III, we demonstrate that the beaconing-based system is able to inform the video-based system about oncoming traffic before any interference on the SCH begins. In Section IV, we show that the video transmission in the SCH can be handled in real time, as well as demonstrating the way to choose the video encoding parameters to maximize the visual quality. Finally, in Section V, we describe the joint beaconing and video-based overtaking assistance system and demonstrate its benefits. Conclusions are presented in Section VI.

II. SYSTEM MODEL

A. Intervehicle Communications

Following the assumptions from [3], we consider a straight two-lane rural road with two platoons of IEEE 802.11p/WAVE-enabled vehicles driving in opposite directions with velocity V and correspondingly approaching each other with a speed of $2V$ (see Fig. 2 and Table I). Each platoon has $N + 1$ vehicles with an intervehicle distance δ . The first vehicle in each of the platoons (i.e., 1 and 3) is referred to as a *leading* vehicle; the rest N are *regular* vehicles (e.g., 2 and 4). The distance between leading vehicles at a given moment in time is denoted by x . Leading vehicles exchange beacons in CCH and transmit the real-time video to regular vehicles in their platoons in the preselected SCHs.

The communication links between the vehicles are described using two signal propagation models.

- 1) *deterministic propagation model*, which is proposed for the analysis of the overtaking assistant in [3];
- 2) *stochastic Nakagami- m propagation model*, which is typically applied in intervehicle communication analysis [9].

Let $P(x)$ be the probability that the packet is successfully received when the distance between vehicles is x .

In the case of the deterministic propagation model from [3], function $P(x)$ is defined as

$$P(x) = \begin{cases} 0.999, & \text{if } x \leq 400 \\ (-0.4x + 210)/100, & \text{if } 400 < x \leq 500 \\ 0.1, & \text{if } 500 < x \leq 600 \\ 0, & \text{if } x > 600. \end{cases} \quad (1)$$

In the Nakagami- m case, the propagation is modeled by the probability density function of the received signal amplitude A , which is defined as

$$f(A, m, \Omega(x)) = \frac{2m^m}{\Gamma(m)\Omega(x)^m} A^{2m-1} \exp\left(-\frac{m}{\Omega(x)} A^2\right) \quad (2)$$

TABLE I
MAIN NOTATIONS

Parameter	Meaning
V	velocity of the platoons
N	number of regular vehicles in a platoon
δ	inter-vehicle distance within a platoon
x	inter-platoon distance
R_b, R_e	beaconing and event-driven messages generation frequencies
R_v, L_v, C_v	video rate, packet length and physical channel rate
$P(x)$	probability of successful packet delivery at the distance x
x_{\max}	maximum distance between the vehicles over which communication between them is possible for the deterministic propagation model
$E[D_n^{(k)}]$	expected end-to-end distortion for slice k in frame n
$p_n^{(k)}$	loss probability for the slice k in frame n
$\Delta_n^{(k)}$	quantization step for the slice k in frame n
$D(\Delta_n^{(k)})$	distortion caused by the quantization with the step $\Delta_n^{(k)}$
D_{loss}	distortion caused by packet losses and error concealment at the decoder
θ, D^0, R^0	operational rate-distortion function model parameters
\hat{D}_{\max}	maximum possible distortion caused by packet loss
FPS	video frame rate
$P_{\text{contact}}(x)$	probability of the event which occurs when the leading vehicle successfully receives the beacon from another leading vehicle first time at distance x
$P_{\text{warning}}(x)$	probability of the event which occurs when all vehicles in the platoon become aware of the contact event at distance x
T_{SI}, T_{SCH}	duration of synchronization and service channel intervals in the WAVE channel switching scheme
T_g	WAVE guard time
W_v	minimum connection window used for the video packet transmission
$AIFS, EIFS, \alpha$	arbitrary and extended inter-frame spaces and $aSlotTime$ in IEEE 802.11p
T_h	duration of physical layer convergence protocol preamble and header
T_s, T_c	duration of successful and collided video packet transmissions
τ	mean video packet transmission time
M	number of packets generated by a leading vehicle during the WAVE synchronization interval
$X(\delta, x)$	mean number of video packets transmitted per SI without collisions by a leading vehicle to a regular vehicle located at a distance δ when the distance between leading vehicles is x

where $\Omega(x)$ defines the average received power at distance x , the value $m \geq 1/2$ identifies the fading intensity and depends on the environment, and Γ is the Gamma function. Following [9], we use a two-ray ground propagation model to calculate $\Omega(x)$, i.e.,

$$\Omega(x) = \begin{cases} \frac{P_t G_t G_r \lambda^2}{(4\pi)^2 x^2}, & \text{if } x \leq x_c \\ \frac{P_t G_t G_r \lambda^2 h_t^2 h_r^2}{x^4}, & \text{if } x > x_c \end{cases} \quad (3)$$

where $x_c = 4\pi h_t h_r / \lambda$ is a crossover distance; (h_t, h_r) and (G_t, G_r) are the antenna heights and antenna gains of the transmitter and receiver, respectively; and λ is a carrier wavelength.

Successful reception of a packet in the absence of interference occurs when the received signal power exceeds the

reception threshold $RxTh$. Therefore, the probability $P(x)$ that the packet is successfully received when the distance between vehicles is x for Nakagami- m model (2) is calculated as

$$P(x) = \int_{RxTh}^{\infty} f(A, m, \Omega(x)) dA. \quad (4)$$

B. End-to-End Video Distortion Model

In the video compression framework, the video frame is represented as a 2-D array of pixels, each of which contains one luma (Y) and two chroma (U and V) color components. Each video frame is separated into nonoverlapping *slices*, which are independently compressed. Each slice is placed in one packet and transmitted over the SCH with video rate R_v .

Compression at the transmitter side includes, at the least, discrete cosine transform (DCT), uniform scalar quantization of DCT coefficients, and entropy coding. Decompression at the receiver side includes decompression of the quantized DCT coefficients by the entropy decoder, dequantization, and inverse DCT.

If the compression and decompression process is considered without any packet losses during wireless transmission, then only the quantization introduces any distortion between captured and reconstructed slices. The choice of the quantization step allows control of the tradeoff between the distortion $D(\Delta_n^{(k)})$ and rate $R_v(\Delta_n^{(k)})$, where $\Delta_n^{(k)}$ is the quantization step for slice with number k in frame n .

The distortion is defined as

$$D(\Delta_n^{(k)}) = \frac{1}{|s_n^{(k)}|} \sum_{(i,j) \in s_n^{(k)}} (y[i, j] - \hat{y}[i, j])^2 \quad (5)$$

where $y[i, j]$ is the value for a sample with coordinates (i, j) in the original slice $s_n^{(k)}$, $\hat{y}[i, j]$ is the corresponding value in the reconstructed slice after dequantization, and $|s_n^{(k)}|$ is the number of samples in the slice. If $\Delta_n^{(k)}$ is small, then, for uniform scalar quantization, we can approximate the distortion as

$$D(\Delta_n^{(k)}) = \frac{(\Delta_n^{(k)})^2}{12}. \quad (6)$$

For the analytical description of the output video bit rate $R_v(\Delta_n^{(k)})$ after the compression with quantization step $\Delta_n^{(k)}$, we use the following operational rate-distortion function model [10]:

$$\begin{aligned} R_v(\Delta_n^{(k)}) &= \frac{\theta}{D(\Delta_n^{(k)}) - D^0} + R^0 \\ &= \frac{\theta}{(\Delta_n^{(k)})^2 / 12 - D^0} + R^0 \end{aligned} \quad (7)$$

where θ , R^0 , and D^0 are the model parameters.

In real systems, distortion can also be caused by packet losses due to channel noise, collision, or packet time-to-live expiry. In this case, video decoders usually use the error concealment algorithm to minimize the visual distortion. (For example, the corresponding slice from the previous reconstructed frame can be displayed, instead of the lost slice.) Then, if the packet was successfully delivered, the video distortion will be equal to $D(\Delta_n^{(k)})$; otherwise, the video distortion will be equal to D_{loss} . Thus, the expected end-to-end distortion [11] is calculated as

$$E[D_n^{(k)}] = (1 - p_n^{(k)}) D(\Delta_n^{(k)}) + p_n^{(k)} E[D_{\text{loss}}] \quad (8)$$

where $p_n^{(k)}$ is the loss probability for the slice $s_n^{(k)}$.

For simplicity, let the video decoder use the following error concealment algorithm: If the slice k from the frame n is not received, then the decoder plays back the slice from the previous reconstructed frame $n - 1$ with the same coordinates. Let us define the distortion in this case as $\hat{D}_{n-1}^{(k)}$. If the corresponding slice in the frame $n - 1$ is also not received, then the decoder plays back the slice from frame $n - 2$ with distortion $\hat{D}_{n-2}^{(k)}$, and so on. Taking into account these assumptions, $E[D_{\text{loss}}]$ can be written as

$$E[D_{\text{loss}}] = (1 - p_{n-1}^{(k)}) \hat{D}_{n-1}^{(k)} + p_{n-1}^{(k)} (1 - p_{n-2}^{(k)}) \hat{D}_{n-2}^{(k)} + \dots \quad (9)$$

In the application considered, video frames are captured from a fast-moving vehicle. Thus, the neighboring frames are not the same, and $E[D_{\text{loss}}]$ can be approximated as

$$E[D_{\text{loss}}] \approx (1 - p_{n-1}^{(k)}) \hat{D}_{n-1}^{(k)} + p_{n-1}^{(k)} \hat{D}_{\text{max}} \quad (10)$$

where \hat{D}_{max} is the maximum possible distortion.

As the visual quality metric, the well-known expected peak signal-to-noise ratio is used, i.e.,

$$E[PSNR] = 10 \lg \frac{y_{\text{max}}^2}{E[D_n^{(k)}]} \quad (11)$$

where y_{max} is the maximum possible value of the luma component. Taking into account that the luma (Y) component is more important for human vision than the chroma (U and V) components, we calculate the expected peak signal-to-noise ratio $E[Y - PSNR]$ for the luma component only.

The calculation process for the end-to-end distortion parameters is shown in Fig. 3. The test video sequence is compressed by the H.264/AVC encoder [12] for all possible values of the quantization steps $\{\Delta_n^{(k)}\}$, which results in the set of video rates $\{R_v(\Delta_n^{(k)})\}$. Then, each video stream is decompressed by the H.264/AVC decoder, and the set of distortions $\{D(\Delta_n^{(k)})\}$ is calculated. The values of the model parameters θ , R^0 , and D^0 are chosen to approximate the operation rate-distortion curve $\{R_v(\Delta_n^{(k)}), D(\Delta_n^{(k)})\}$ by function (7). The value of $\hat{D}_{n-1}^{(k)}$ is calculated as the average distortion between the frames with numbers n and $n - 1$ in the test video sequence. Finally, \hat{D}_{max} is an analogous value for frames n and $n - 8$.

We use JM codec [13] as the reference software for the H.264/AVC standard. The test video sequence ‘‘Highway’’ [14]

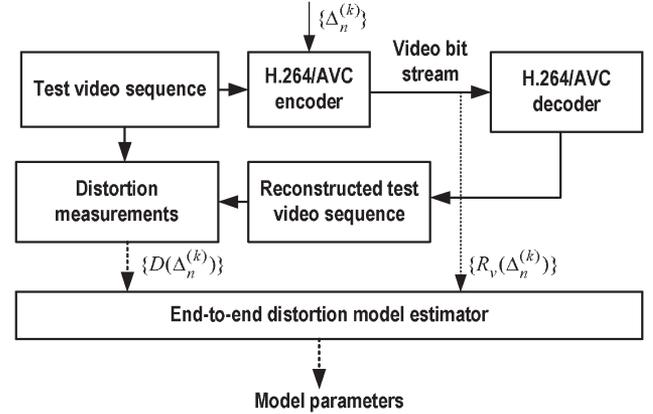


Fig. 3. Estimation of the end-to-end video distortion model parameters.

with frame resolution 352×288 , which contains captured video from a moving vehicle on the road, is used in our experiments. For such a case, $\theta = 9735$, $R_0 = 15.5$, $D_0 = 0$, $\hat{D}_{n-1}^{(k)} = 100$, and $\hat{D}_{\text{max}} = 500$.

III. PERFORMANCE EVALUATION OF THE BEACON-BASED SYSTEM

A. Probability Distributions of Contact and Warning Events

Let us consider in detail the operation of the beacon-based system. Both the leading and regular vehicles generate beacons with a frequency of R_b Hz. As soon as the beacon from one leading vehicle is received by another leading vehicle, the event-driven message is generated to warn all the regular vehicles in the corresponding platoon. The event-driven message is broadcast with a frequency of R_e Hz. For simplicity of the mathematical analysis, we assume that there is a one-hop connectivity between the leading vehicle and all the regular vehicles in the platoon. This assumption does not hold for the long platoons. In such a case, fast dissemination of the event-driven message to the distant regular vehicles can be easily provided by adopting, for example, the algorithms from [9].

Now, let us define two events that are the most important for the operation of the beacon-based system.

- 1) *Contact*, which occurs when the leading vehicle successfully receives the beacon from another leading vehicle for the first time. When contact occurs, the leading vehicle becomes aware of the approaching of oncoming traffic and then starts to disseminate the warning message to warn all the other vehicles in the platoon. Moreover, this information is also used to inform the video rate control about the changes in the SCH load.
- 2) *Warning*, which occurs when all vehicles in the platoon become aware about the contact event. On receipt of the warning message, the information about the oncoming traffic is provided to the driver of the regular vehicle. We assume that the warning event happens when the *last* vehicle in the platoon, which experiences the worst reception conditions, receives the event-driven warning message.

Probability distributions of these events, with respect to the interplatoon distance, are key performance metrics of the beacon-based overtaking assistance system.

In our analysis, we assume that the event-driven messages are transmitted with the highest EDCA priority and beacons are transmitted with the lowest EDCA priority. Due to the relationships between the minimum contention windows and arbitrary interframe spaces in both of the priority classes, this guarantees that the beacons and the event-driven messages sent by vehicles in the platoon never collide [1]. Moreover, the collision probability between the beacons of the leading vehicles is negligible.

Let x_{\max} be the maximum distance between the leading vehicles when the communication between them is possible. Then, the number of beacons that were transmitted by the leading vehicle when distance $x \leq x_{\max}$ can be calculated as $n_b(x) = \lfloor ((x_{\max} - x)/2V)R_b \rfloor$.

Then, the probability that a contact event occurs until distance x is

$$P_{\text{contact}}(x) = 1 - \prod_{i=0}^{n_b(x)-1} (1 - P(l_i)) \quad (12)$$

where $l_i = x_{\max} - (i/R_b)2V$. Let the probability that contact happens exactly at distance l_i be

$$\hat{P}_{\text{contact}}(l_i) = P(l_i) \prod_{k=0}^{i-1} (1 - P(l_k)) \quad (13)$$

and if k , which is the number of transmission attempts, has been performed by the leading vehicle, the probability that the last vehicle in the platoon receives the warning message will be

$$\hat{P}_{\text{warning}}(k) = (1 - (1 - P(N\delta))^k). \quad (14)$$

It is assumed that rebroadcasting of the event-driven messages by the leading vehicle is continued until all the regular vehicles in its platoon successfully receive it. Then, the probability that a warning event occurs within distance x is

$$P_{\text{warning}}(x) = \sum_{i=0}^{n_b(x)-1} \hat{P}_{\text{contact}}(l_i) \hat{P}_{\text{warning}}(n_b(x) - i). \quad (15)$$

B. Discussion on the Performance of Beaconing-Based System

We use practical sets of parameter values from [3]: $V = 90$ km/h, $\delta = 50$ m, and $R_b = R_e = 20$ Hz. Propagation and transceiver parameters are given as follows [9]: $P_t = 12.95$ dBm, $h_t = h_r = 1.5$ m, $G_t = G_r = 2.512$ dBm, $\lambda = 50.85$ mm, and $RxTh = -92$ dBm.

The probability distributions of contact and warning events with respect to distance are shown in Fig. 4 for a deterministic propagation model and in Figs. 5 and 6 for the Nakagami- m model for $m = 1$ (low fading) and $m = 7$ (severe fading).

The following conclusions can be drawn:

- 1) *Contact* event occurs at distances of about 500 m with a probability close to one for the deterministic propagation

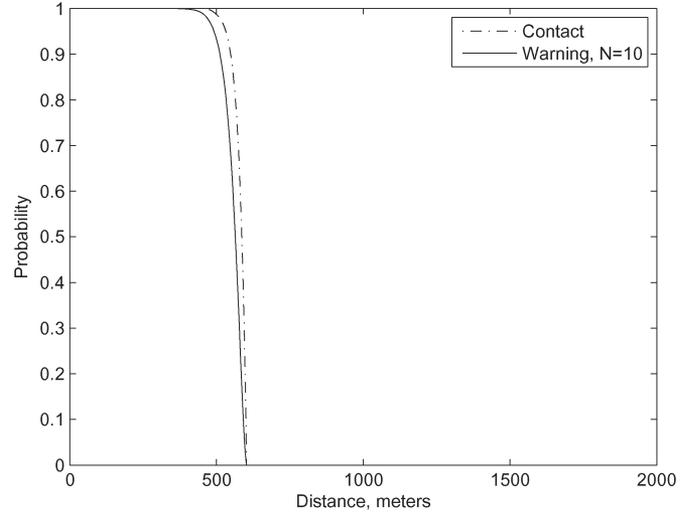


Fig. 4. Probabilities of “Contact” and “Warning” events, depending on distance for the deterministic propagation model.

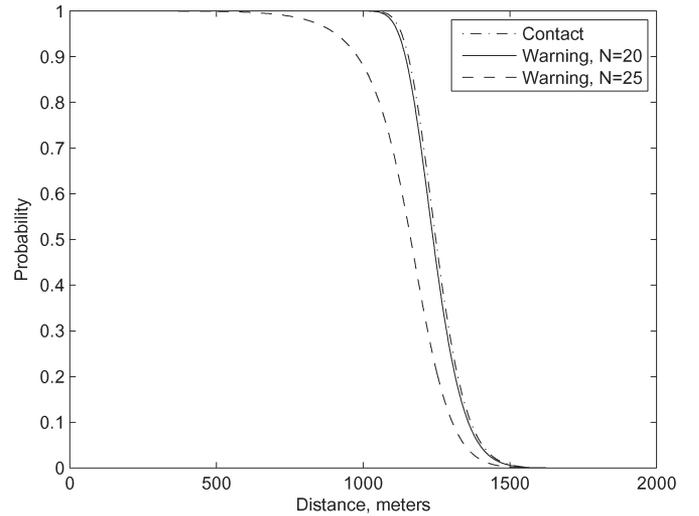


Fig. 5. Probabilities of “Contact” and “Warning” events, depending on distance for the Nakagami- m model with fading intensity $m = 1$.

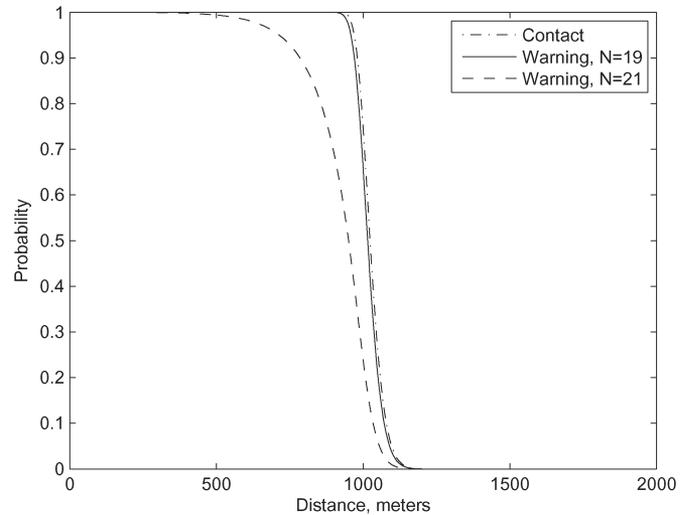


Fig. 6. Probabilities of “Contact” and “Warning” events, depending on distance for the Nakagami- m model with fading intensity $m = 7$.

model. Hence, information about the change of SCH channel load becomes available at the video bit rate controller, when the influence of the upcoming platoon on the reception of the video stream is almost negligible. Analogous distances for the Nakagami- m model are on the order of 1000 m for both the fading intensities considered. Thus, the video bit rate controller is informed about the upcoming platoon well in advance, i.e., not less than 20 s before the actual platoon meeting.

- 2) *Warning* to all the drivers in the platoon will be issued by the beacon-based system almost at the same time as the contact event occurs in the case of the deterministic propagation model. For the intervehicle distance considered, a maximum of ten regular vehicles can be warned. In this case, the last vehicle will get a warning at around 500 m before the meeting of the platoons. As for the Nakagami- m model, one can see that independently of the chosen fading intensity, 20 vehicles can be notified with a probability close to one at a distance of 1000 m. The beacon-based system thus provides a timely warning to the driver.

IV. PERFORMANCE EVALUATION OF THE VIDEO-BASED SYSTEM

A. End-to-End Video Transmission Latency

The latency ΔT between the time when some slice has been captured and the time when this slice is shown by the receiver device display consists of the following components:

$$\Delta T = \Delta T_e + \Delta T_{tb} + \Delta T_c + \Delta T_{rb} + \Delta T_d \quad (16)$$

where ΔT_e is the encoding latency, ΔT_d is the decoding latency, ΔT_{tb} is the transmitter buffering latency, ΔT_{rb} is the receiver buffering latency, and ΔT_c is the channel transmission latency.

Let us suppose that the encoder and the decoder operate in real time. Then, taking into account the fact that all packets have equal size and that, in the boundary case, one packet can contain a whole video frame, the values ΔT_e and ΔT_d do not exceed $1/FPS$, where FPS is the video frame rate.

To control the transmitter buffering latency, we propose using the following transmission rule: All video packets related to the frames captured during a given WAVE synchronization interval (SI) [2] are transmitted in the SCH interval of this SI. All the video packets still in the transmitter buffer at the end of the SI are dropped. Therefore, $\Delta T_e + \Delta T_{tb} + \Delta T_c \leq T_{SI}$, where $T_{SI} = 100$ ms is the duration of the SI.

Assuming that a decoding process starts immediately after the reception of a video packet, then ΔT_{rb} is not significant, and the end-to-end latency is given by

$$\Delta T \leq T_{SI} + \frac{1}{FPS}. \quad (17)$$

For the typical frame rate $FPS = 30$, the end-to-end latency is $\Delta T \leq 133$ ms, which is acceptable for the application being studied.

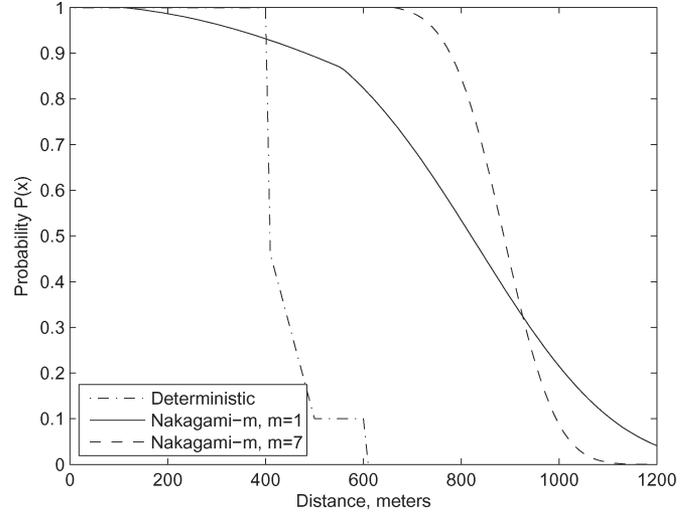


Fig. 7. Probabilities of successful packet delivery, depending on distance for the deterministic and the Nakagami- m propagation models.

B. SCH Channel Throughput for Deterministic Model

For simplicity of mathematical analysis, we assume that the channel conditions do not vary during the SI. As shown in Fig. 7, since the duration of the SI is 100 ms, the interplatoon distance during the SI is only decreased by 5 m, which results in an insignificant change in the probability of successful reception $P(x)$. This assumption is therefore realistic within the framework of the propagation models considered.

Let us consider the transmission of video packets from vehicle 1 to vehicle 2 and from vehicle 3 to vehicle 4 (see Fig. 1). According to the EDCA protocol rules [1], the average time needed to transmit every broadcast video packet is equal to

$$\tau = \frac{W_v - 1}{2} \sigma + T_s \quad (18)$$

where $T_s = T_h + L_v/C_v + AIFS$; $L_v = 800$ bytes is the packet size; $C_v = 6$ Mb/s is the channel rate; $T_h = 40$ μ s is the duration of the PHY convergence protocol preamble and header; $W_v = 16$ is the minimum contention window; $AIFS = 64$ μ s is an arbitrary interframe space; and $\sigma = 16$ μ s is an *aSlotTime* defined by [1].

Transmissions from vehicle 3 influence the decisions of vehicle 1 to access the SCH channel. Assuming an approach similar to [15], we find that, for the saturation conditions, the SCH is divided into “virtual” slots of different durations, and vehicles 1 and 3 are both accessing the SCH in every slot with a constant probability $\rho = 2/(W_v + 1)$. The mean duration of the virtual slot can thus be computed as follows:

$$T_{vs}(\delta, x) = (1 - \rho)^2 \sigma + \rho T_s + (1 - \rho) \rho P(x) T_s + (1 - \rho) \rho (1 - P(x)) T_c \quad (19)$$

where $T_c = T_h + L_v/C_v + EIFS$ and $EIFS = 91$ μ s is an extended interframe space. The aforementioned equation incorporates the probabilities and the mean durations of four events, which are possible in SCH.

- 1) Vehicles 1 and 3 do not transmit in the same virtual slot.

- 2) Vehicle 1 broadcasts in some virtual slot.
- 3) Vehicle 1 observes successful transmission from vehicle 3.
- 4) Vehicle 1 observes unsuccessful transmission from vehicle 3.

The useful duration of the SCH interval can be found by applying the reasoning from [16]: $T = T_{\text{SCH}} - T_g - T_h - L_v/C_v$, where $T_g = 4$ ms is the WAVE guard interval, and $T_{\text{SCH}} = 50$ ms is the duration of the SCH interval. Then, the average number of virtual slots in each SCH is T/T_{vs} , and the mean number of video packets transmitted in the SI from vehicle 1 to vehicle 2 without collisions is

$$X(\delta, x) = \begin{cases} \lfloor \frac{T}{\tau} \rfloor, & \text{if } x > 600 \\ \rho(1 - \rho) \lfloor \frac{T}{T_{\text{vs}}(\delta, x)} \rfloor, & \text{if } x \leq 600. \end{cases} \quad (20)$$

C. SCH Channel Throughput for Nakagami- m Model

We model the vehicles air interfaces as in [9]. Three power thresholds are considered.

- 1) *Carrier-sense threshold* ($CSTh = -96$ dBm): If the received signal power is lower than $CSTh$, then the packet is assumed to be dropped without any influence on the receiver; otherwise, the signal is sensed and prevents an interface from transmission.
- 2) *Reception threshold* ($RxTh = -92$ dBm): If there only one transmission and received signal power is more than $RxTh$, then the packet is successfully received.
- 3) *Capture threshold* ($CpTh = 7$ dB): If there are two concurrent transmissions, then the packet is successfully received only if its power is larger than $RxTh$, and the difference between this power and the power of the interfering packet exceeds $CpTh$.

Receiving signal powers from vehicle 1 and from vehicle 3 at vehicle 2 are denoted as η_1 and η_3 , respectively; then, the following events in SCH occur (see Figs. 8 and 9):

- 1) *Event A*: Vehicle 1 and vehicle 3 simultaneously transmit; herewith, $\eta_1 \geq RxTh$, $\eta_3 \geq CSTh$, and $\eta_1 - \eta_3 \leq CpTh$. This event will be treated at vehicle 2 as a collision. For simplification, we assume that $\eta_1 = \Omega(\delta)$ and $\eta_1 - CpTh > CSTh$. The probability of this event is then

$$P_A(\delta, x) \approx \int_{\Omega(\delta) - CpTh}^{\infty} f(A, m, \Omega(x + \delta)) dA. \quad (21)$$

- 2) *Event B*: Vehicle 3 transmits and vehicle 1 senses the SCH; herewith, $CSTh \leq \eta_3 < RxTh$. This event corresponds to the case when vehicle 1 senses some transmission in the SCH and treats it as a collision. The probability of this event is then

$$P_B(x) = \int_{CSTh}^{RxTh} f(A, m, \Omega(x)) dA. \quad (22)$$

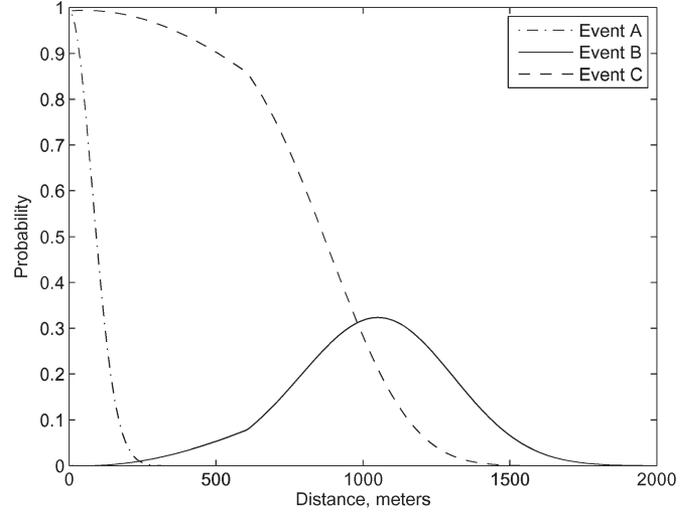


Fig. 8. Probabilities of events A, B, and C, depending on distance for the Nakagami- m model with fading intensity $m = 1$.

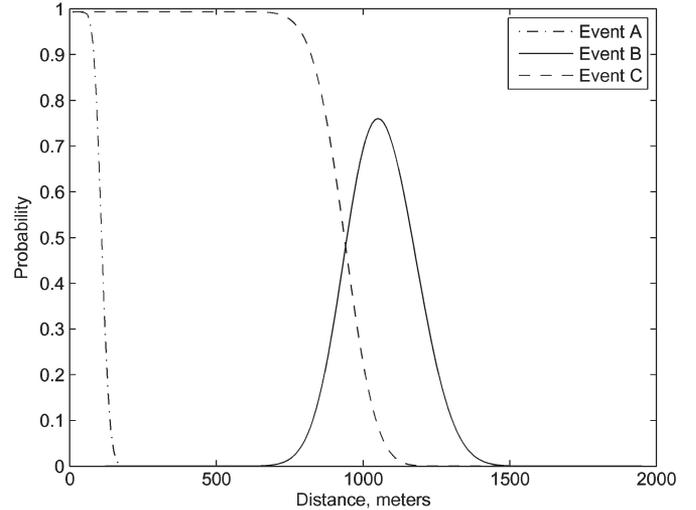


Fig. 9. Probabilities of events A, B, and C, depending on distance for the Nakagami- m model with fading intensity $m = 7$.

- 3) *Event C*: Vehicle 3 transmits and vehicle 1 senses the SCH; herewith, $\eta_3 \geq RxTh$. This event corresponds to the case where vehicle 1 senses some transmission in the SCH and treats it as a success. The probability of this event is then

$$P_C(x) = \int_{RxTh}^{\infty} f(A, m, \Omega(x)) dA. \quad (23)$$

The mean duration of the virtual slot can be analogously computed to (19) as follows:

$$T_{\text{vs}}(\delta, x) = (1 - \rho)^2 \sigma + \rho T_s + (1 - \rho) \rho P_B(x) T_c + (1 - \rho) \rho P_C(x) T_s + (1 - \rho) \rho (1 - P_B(x)) \sigma. \quad (24)$$

Then, the average number of virtual slots per the SCH is T/T_{vs} , and the mean number of video packets transmitted in

the SI from vehicle 1 to vehicle 2 without collisions is

$$X(\delta, x) = (\rho^2 (1 - P_A(\delta, x)) + \rho(1 - \rho)) \frac{T}{T_{vs}(\delta, x)} \quad (25)$$

where the first term corresponds to the case, where vehicles 1 and 3 simultaneously transmit, but the packet from vehicle 1 is successfully received by vehicle 2; on the other hand, the second term corresponds to the case where only vehicle 1 transmits.

D. Quantization Step Selection

Let us assume that the video camera captures FPS frames per second and n_0 is the number of the frame, which was captured by the camera of the leading vehicle 1 at the moment when the distance between the leading vehicles x_{\max} .

Given that the video bit rate is $R_v(\Delta_n^{(k)})$, the number of packets generated by a leading vehicle during WAVE SI can be computed as

$$M = \left\lceil \frac{R_v(\Delta_n^{(k)}) T_{SI}}{L_v} \right\rceil. \quad (26)$$

The packet loss probability for slice k in frame n due to the channel conditions can then be calculated as

$$p_n^{(k)} = 1 - \frac{\min(X(\delta, x), M)}{M} P(\delta) \quad (27)$$

where $P(\delta)$ is the probability that the packet from vehicle 1 is successfully received by vehicle 2 in the absence of interference from vehicle 3 (see Section III).

As was shown in Section IV-B, the channel slowly changes during SI; therefore, we can assume that $p_n^{(k)} \approx p_{n-1}^{(k)}$. Thus, the minimum of the expected end-to-end distortion (8) for the given quantization step $\Delta_n^{(k)}$ is achieved if

$$p_n^{(k)} = \frac{(\Delta_n^{(k)})^2 / 12 - \hat{D}_{n-1}^{(k)}}{2(\hat{D}_{\max} - \hat{D}_{n-1}^{(k)})}. \quad (28)$$

As was noted in Section II-B, we use small quantization steps $\Delta_n^{(k)}$, and neighboring frames are not similar ($\Delta^2/12 - \hat{D}_{n-1}^{(k)} < 0$), which leads to the conclusion that the expected distortion is a nondecreasing function of the interval $p_n^{(k)} \in [0, 1]$. Thus, from (27), the minimum expected distortion is achieved for the minimal possible value of $p_n^{(k)}$ or when $M = X(\delta, x)$. From (7) and (26), it follows that this equality is achieved when

$$\Delta_n^{(k)} = \sqrt{\frac{12\theta T_{SI}}{X(\delta, x)L_v - R^0 T_{SI}} + 12D^0} \quad (29)$$

where $x = x_{\max} - (2V(n - n_0))/FPS$.

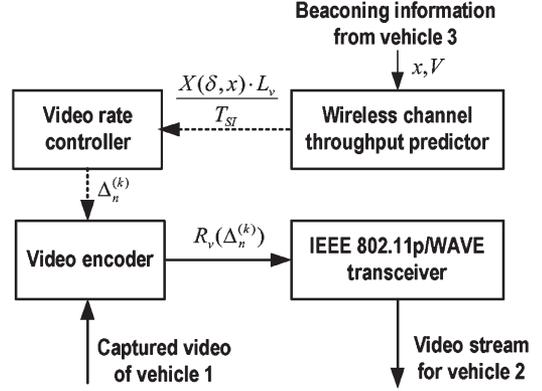


Fig. 10. General structure of the joint overtaking system.

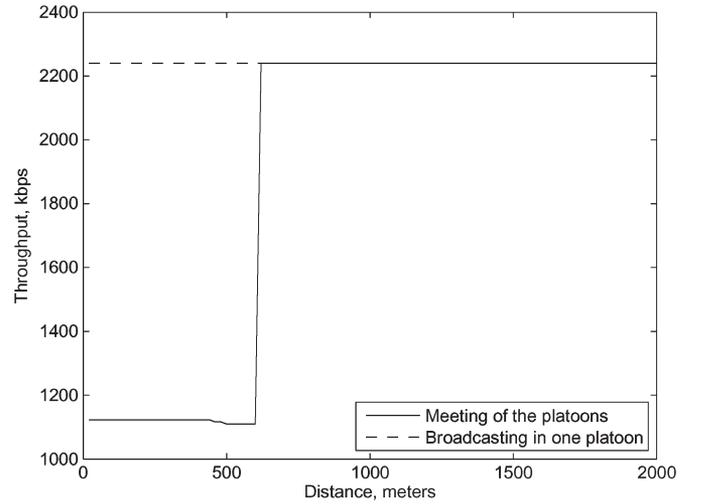


Fig. 11. Multiple-access channel throughput, depending on distance for the deterministic propagation model.

V. EVALUATION OF A JOINT BEACONING AND VIDEO-BASED SYSTEM

Operation of the proposed joint beaconing and video-based system is shown in Fig. 10 for the typical case shown in Fig. 1. It consists of the following steps:

- 1) Vehicle 1 receives from vehicle 3 the beaconing information, which contains its velocity V and current distance x . This information is used by vehicle 1 to predict the changes of the SCH throughput during the approach of the platoons, as explained in Sections IV-B and C. From Figs. 11–13, one can see that the SCH throughput rapidly decreases with the decrease in distance x . Independently of the propagation model, the SCH throughput of the leading vehicle is decreased by a factor of two, which makes it important to adapt the video rate according to the channel load. As illustrated in Section III, this can be done well in advance by the use of beaconing information.
- 2) The video rate controller selects quantization step $\Delta_n^{(k)}$ for each slice to minimize the expected end-to-end distortion, as discussed in Section IV-D. The influence of the quantization step on end-to-end distortion is shown in

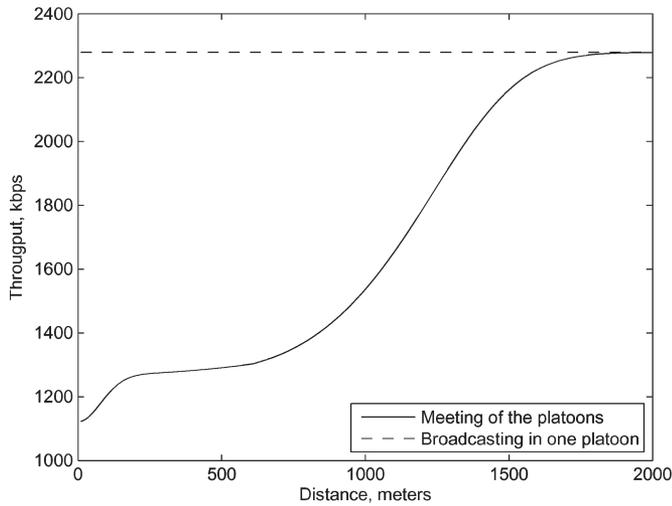


Fig. 12. Multiple-access channel throughput, depending on distance for the Nakagami- m model with fading intensity $m = 1$.

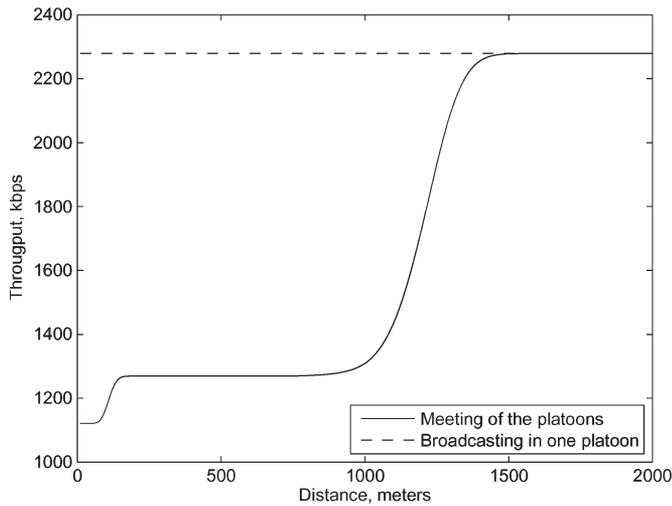


Fig. 13. Multiple-access channel throughput, depending on distance for the Nakagami- m model with fading intensity $m = 7$.

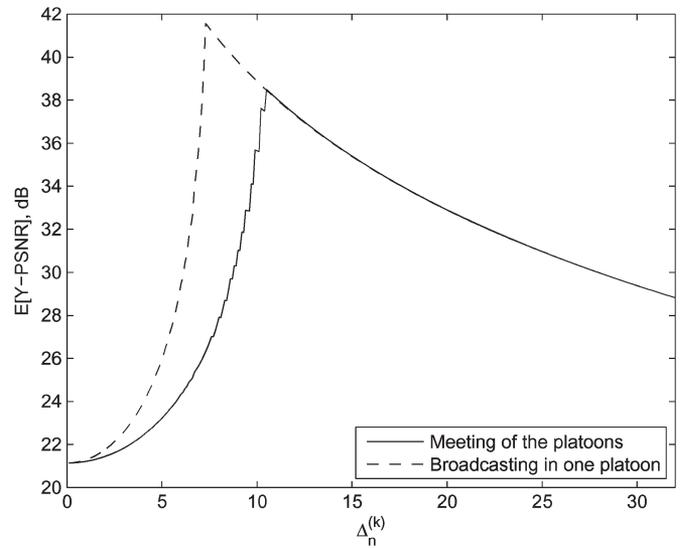


Fig. 14. Expected video quality, depending on quantization step for the deterministic propagation model.

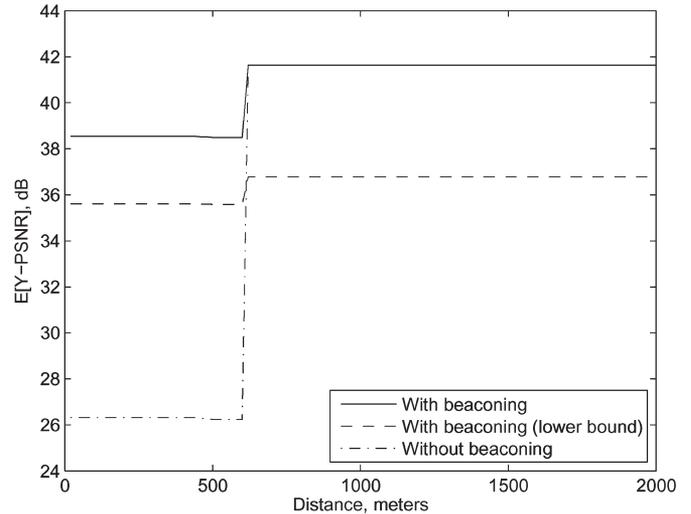


Fig. 15. Expected visual quality with and without use of beaconing information for the deterministic propagation model.

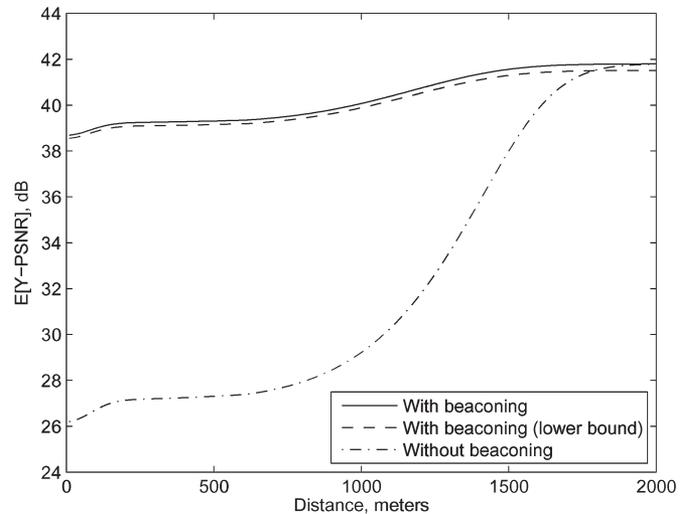


Fig. 16. Expected visual quality with and without use of beaconing information for the Nakagami- m model with fading intensity $m = 1$.

Fig. 14. One can see that improper choice of the quantization step leads to serious visual quality degradation.

- 3) Video information captured by vehicle 1 is compressed by video encoder using the selected quantization step and transmitted to vehicle 2, where it is displayed to the driver. The visual quality in vehicle 2, for the cases when beacons from vehicle 3 are successfully received by vehicle 1, as well as without beaconing information, are shown in Figs. 15–17. One can see that when beaconing information is not used, the video quality becomes unacceptably low. If the beacons are available and the video rate is adapted according to the channel throughput early enough, then this guarantees acceptable quality during the overall time period of the approach of the platoons.

Since the proposed system concerns the active automotive safety, it is extremely important to guarantee real-time video information delivery with acceptable visual quality. First, as it is demonstrated in Section IV-A, the estimated upper bound for

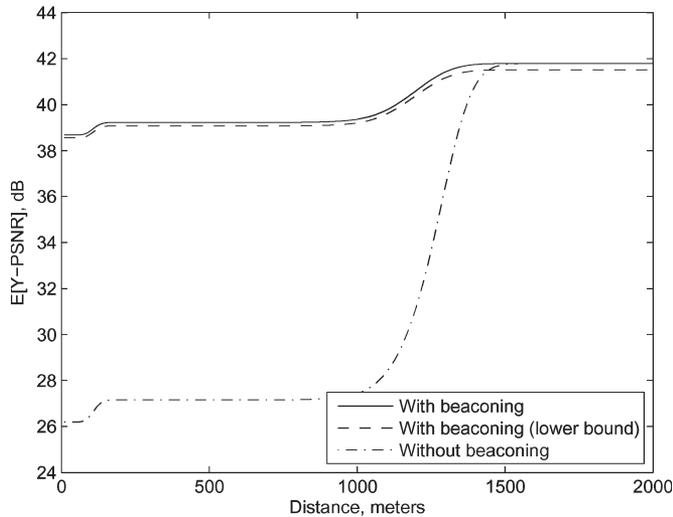


Fig. 17. Expected video quality with and without using of beaconing information for the Nakagami-m model with fading intensity $m = 7$.

the end-to-end latency of video transmission is 133 ms, which is adequate for the driver's timely reaction. Second, we have provided the lower 99.7%-confidence bound for the expected $PSNR$ to estimate the worst possible visual quality during the system operation (see Figs. 15–17). The results show that for the frame resolution 352×288 and a frame rate of 30 frame/s, the lower bound of the expected $PSNR$ is about 36 dB, which corresponds to the reasonable quality level. One can also see that, when beaconing information is not used, then the $PSNR$ value decreases by up to 25 dB, which is unacceptably low for the overtaking scenario.

We, therefore, conclude that our proposed joint beaconing-based and real-time video-based cooperative overtaking assistance system is functionally operative for the applications considered and that it can be implemented in practice.

VI. CONCLUSION

Up-to-date proposals for cooperative overtaking assistance systems rely on either the beaconing of status messages or real-time video transmission. The beaconing-based system provides limited information to the driver about the upcoming traffic in the form of warnings. Introduction of the video-based system can extend the amount of driver information, thus increasing the reliability of the assistance. However, when two platoons of vehicles are meeting each other, the quality of the video information can undergo significant degradation since a common multiple-access communication channel is used. Our proposal gives the guarantee of low latency and acceptable visual quality by making use of the additional information obtained from the beaconing.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers and the associate editor for their time and valuable constructive feedback.

REFERENCES

- [1] *Wireless Access in Vehicular Environments*, IEEE Std. 802.11p, Jul. 2010.
- [2] *IEEE Standard for Wireless Access in Vehicular Environments (WAVE)—Multi-Channel Operation*, IEEE Std. 1609.4-2010, Feb. 2011.
- [3] A. Böhm, M. Jonsson, and E. Uhlemann, "Adaptive cooperative awareness messaging for enhanced overtaking assistance on rural roads," in *Proc. 4th Int. Symp. WIVEC*, San Francisco, CA, Sep. 2011, pp. 1–5.
- [4] N. N. Qadri, M. Fleury, M. Altaf, and M. Ghanbari, "Multi-source video streaming in a wireless vehicular ad hoc network," *IET Commun.*, vol. 4, no. 11, pp. 1300–1311, Jul. 2010.
- [5] M. Asefi, J. W. Mark, and X. Shen, "An application-centric inter-vehicle routing protocol for video streaming over multi-hop urban VANETs," in *Proc. IEEE ICC*, Kyoto, Japan, Jun. 2011, pp. 1–5.
- [6] L. Zhou, Y. Zhang, K. Song, W. Jing, and A. Vasilakos, "Distributed media services in P2P-based vehicular networks," *IEEE Trans. Veh. Technol.*, vol. 60, no. 2, pp. 692–703, Feb. 2011.
- [7] F. Soldo, C. Casetti, C.-F. Chiasserini, and P. A. Chaparro, "Video streaming distribution in VANETs," *IEEE Trans. Parallel Distrib. Syst.*, vol. 22, no. 7, pp. 1085–1091, Jul. 2011.
- [8] C. Olaverri-Monreal, P. Gomes, R. Fernandes, F. Vieira, and M. Ferreira, "The see-through system: A VANET-enabled assistant for overtaking maneuvers," in *Proc. IEEE IV*, San Diego, CA, Jun. 2010, pp. 123–128.
- [9] M. Torrent-Moreno, J. Mittag, P. Santi, and H. Hartenstein, "Vehicle-to-vehicle communication: Fair transmit power control for safety-critical information," *IEEE Trans. Veh. Technol.*, vol. 58, no. 7, pp. 3684–3707, Sep. 2009.
- [10] K. Stuhlmüller, N. Farber, M. Link, and B. Girod, "Analysis of video transmission over lossy channels," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 6, pp. 1012–1032, Jun. 2000.
- [11] F. Zhai, Y. Eisenberg, T. N. Pappas, R. Berry, and A. K. Katsaggelos, "Rate-distortion optimized hybrid error control for real-time packetized video transmission," *IEEE Trans. Image Process.*, vol. 15, no. 1, pp. 40–53, Jan. 2006.
- [12] "Advanced video coding for generic audiovisual services," ITU-T Recommendation H.264 and ISO/IEC 14496-10 (AVC), 2009.
- [13] H.264/AVC JM Reference Software. [Online]. Available: <http://iphome.hhi.de/suehring/tml/>
- [14] Xiph.org Test Media. [Online]. Available: <http://media.xiph.org/video/derf/>
- [15] G. Bianchi, "Performance analysis of the IEEE 802.11 distributed coordination function," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 3, pp. 535–547, Mar. 2000.
- [16] C. Campolo, A. Vinel, A. Molinaro, and Y. Koucheryavy, "Modeling broadcasting in IEEE 802.11p/WAVE vehicular networks," *IEEE Commun. Lett.*, vol. 15, no. 2, pp. 199–201, Feb. 2011.



Alexey Vinel (M'07) received the B.S. (Hons.) and M.S. (Hons.) degrees in information systems from Saint Petersburg State University of Aerospace Instrumentation, St. Petersburg, Russia, in 2003 and 2005, respectively, and the Ph.D. (candidate of science) degree in technical sciences from the Institute for Information Transmission Problems, Russian Academy of Sciences, Moscow, Russia, in 2007.

He is currently a Researcher with the Department of Communications Engineering, Tampere University of Technology, Tampere, Finland. His research interests include multiple-access protocols and intelligent transportation systems.

Dr. Vinel has been a Fellow of the Alexander von Humboldt Foundation since 2008 and a member of organizing and technical committees of many international conferences.



Evgeny Belyaev received the Engineer degree in automated systems of information processing and control and the Ph.D. (candidate of science) degree in technical sciences from the State University of Aerospace Instrumentation, St. Petersburg, Russia, in 2005 and 2009, respectively.

He is currently a Researcher with the Institute of Signal Processing, Tampere University of Technology, Tampere, Finland. His research interests include real-time video compression and transmission, video source rate control, scalable video coding, motion estimation, and arithmetic encoding.



Karen Egjazarian (SM'96) was born in Yerevan, Armenia, in 1959. He received the M.Sc. degree in mathematics from Yerevan State University in 1981, the Ph.D. degree in physics and mathematics from Moscow State University, Moscow, Russia, in 1986, and the D.Tech. degree from Tampere University of Technology (TUT), Tampere, Finland, in 1994.

He has been a Senior Researcher with the Department of Digital Signal Processing, Institute of Information Problems and Automation, National Academy of Sciences of Armenia, Yerevan. Since

1996, he has been an Assistant Professor with the Institute of Signal Processing, TUT, where he is currently a Professor, leading the Transforms and Spectral Methods Group. His research interests are applied mathematics, signal processing, and digital logic.



Yevgeni Koucheryavy (SM'09) received the Ph.D. degree from Tampere University of Technology (TUT), Tampere, Finland, in 2004.

Prior joining TUT, he spent five years in the industry with R&D LONIIS, St. Petersburg, Russia, where he held various technical and managerial positions. He is an invited expert for ITU-T and the Skolkovo Foundation (Russia) and acts as an External Reviewer for the state funding agencies of several European countries. He is currently a Professor with the Department of Communications Engineering, TUT.

His current research interests include heterogeneous wireless communications and systems, network and services performance evaluation, Internet of Things, and machine-to-machine communications.