Realizable Video Transmission Over Fading Channels with Full Channel State Information

Slavche Pejoski and Venceslav Kafedziski

Abstract—**A novel method for wireless video transmission is presented, where channel coherence interval is shorter than the video frame duration and causal channel state information (CSI) is available at the transmitter. Triplets of bit error probability, outage probability and obtainable bits for given outage probability are calculated. Based on these triplets, the video sequence is optimally encoded at the beginning of the video frame. During transmission, parameters at physical layer are chosen to provide bit error probability lower than predetermined one.**

Keywords—**video transmission, fading channel, causal channel state information, packet error probability**

I. Introduction

The real time video transmission has the potential of becoming leading service in the world of telecommunications and it receives a lot of attention.

Real time video transmission is characterized by two most important features: low tolerable delay and variable bit rate. The wireless channel has its own features, among which most important are the restricted resources (bandwidth and power) and the time varying nature of the channel gain. With the continuous advance in telecommunications, there is a possibility to estimate the channel state (channel gain) and feed back this information to the transmitter. This channel state information can be used to choose the optimal modes for different parameters of source and channel encoders. A number of papers [1], [2] [3] show that joint source and channel coding can improve the quality of the received video. This cross-layer paradigm is easily applicable when the channel is known and constant for the whole duration of the video frame transmission and is thoroughly explained in [1] and [2]. In situations when the channel coherence interval (period with equal channel gain) is shorter than the duration of the video frame, it is impossible to know all the channel states that correspond to given video frame at the time instant of encoding that video frame.

In the literature, three approaches are known to be useful in this setting. The first approach employs scalable encoders, for example FGS (Fine Granularity Scalability) video encoders. They encode the video frame using several layers. The adaptation process in this system consists of adapting the physical layer to guarantee some minimal bit error rate (BER) (usually 10^{-6}) and subsequently sending the base layer and part of the enhancement layer, based on their contribution to the overall quality of the decoded video. The down side of this approach is the well known coding inefficiency when low delay is needed. This approach can be found in [4].

The second approach considers the channel as unknown to transmitter and uses coding over subchannels. This approach is quite complicated because it requires interleaving that is adapted to the channel and is usually done with binary modulation. The second disadvantage of this method comes from not using channel information at the transmitter. This approach can be found in [3], [5].

In the third approach the adaptation is done in parts i.e. the video is divided into a number of parts equal to the number of channel coherence intervals corresponding to transmission of a single video frame. So in a given coherence interval specific part of the video and the channel are jointly encoded. The basics of this approach can be found in [1].

In our previous work [6], we investigated the impact of different parameters at the physical layer on the received video quality and we found that variable power has lowest impact. In [6] we used a system with full channel knowledge for all channel states during the video frame transmission. Here we extend the work in [6] by using the conclusions drawn there and making the system causal i.e. the transmitter knows the CSI about the current channel state. In the method several values of outage probability are chosen. Then, based on the statistics of the channel, different BERs are obtained. Each BER corresponds to specific outage probability and for each BER we assume that the packet error rate for video packet with average length, due to outage, is equal to the packet error rate when the system is not in outage. Having the pair of BER and outage probability, the maximal number of bits that can be used in the system at outage probability lower then the calculated outage probability is found. To this number of bits additional bits corresponding to the first channel coherence interval (in this coherence interval the outage probability is zero, but the calculated BER needs to be guaranteed) are added. Based on the triplet (BER, outage probability, number of bits) the source and channel coding are performed. The expected distortion computed at the transmitter is used as the utility function. During transmission, the modes (made of the physical layer parameters) are set to guarantee BER lower than the chosen one. If there are not enough resources to send all the bits, the bits that cannot be sent are discarded.

The paper consists of four sections. In the second section we give description of the system and the proposed algorithm. In the third section we evaluate the proposed algorithm by simulation and in the last section we give conclusion.

S. Pejoski is with the Faculty of Electrical Engineering and Information Technologies, University Cyril and Methodius, Skopje, Republic of Macedonia e-mail: slavchep@feit.ukim.edu.mk.

V. Kafedziski is with the Faculty of Electrical Engineering and Information Technologies, University Cyril and Methodius, Skopje, Republic of Macedonia e-mail: kafedzi@feit.ukim.edu.mk.

II. System and Algorithm Description

The system consists of two entities, one sending the video and the other receiving it. In our system we allow the transmitter to change the parameters of the physical and application layer. Based on the conclusions in [6], we allow the transmitter to choose modes for modulation and coding rate at the physical layer. We use "m" to denote the chosen modulation type, from the set M of all available modulation types and "c" to denote the coding rate, chosen from the set C of all available coding rates. Every combination of "m" and "c" results in mode "v" from the set V of all available combinations. At the application layer the coder can choose the modes of source coding for quantization and coding mode (INTRA/INTER). The chosen mode at the application layer is marked with μ and is chosen from the set of all modes β. The sender starts from the video sequence of raw video frames and encodes it in an adaptive manner in order to achieve best quality at the receiver. During the adaptation process, in order to account for the losses in the wireless channel and the discrepancy between reference frames at the receiver and transmitter, the expected distortion at the decoder is calculated at the encoder using:

$$E[D_n^j] = E[(f_n^j - \tilde{f}_n^j)^2] \tag{1}$$

In (1) f_n^j is the value of the j-th pixel in the n-th video frame, and $\tilde{f}_n^j$ is the value of the j-th pixel of the n-th video frame at the receiver. The calculation of the expected distortion can be done in a recursive manner using the algorithms described in [7] or [8]. The expected distortion depends on the packet error probability, that can be calculated using:

$$P_p = P_{out} + (1 - P_{out})(1 - (1 - P_b)^L). \tag{2}$$

In (2) P_{out} is the outage probability in the system i.e. the probability that the number of bits, used by the video encoder, is larger than the one obtained in the system that guarantees BER lower than P_b. L is the number of bits in the packet.

When coding rate is $r_c = 1$, the BER for binary modulation and MQAM can be respectively calculated from:

$$P_b = Q(\sqrt{2\gamma}) \tag{3}$$

and

$$P_b = 4\frac{\sqrt{M}-1}{\sqrt{M}\cdot\log_2 M}Q(\sqrt{\frac{3\gamma}{M-1}}) \tag{4}$$

If the coding rate is lower than 1, the equations will change in accordance with the channel coding used in the system. In cases where convolutional coding is used, equations for BER calculations can be found in [9].

The optimization process starts at the beginning of the first coherence interval in the video frame transmission interval. The process finds triplets of BER, outage probability and number of bits that can be used to encode the video frame at the specified outage probability. In order to explain the relationship between the elements of the triplet, we will take a look at their calculation for system where the coding rate is $r_c = 1$ and three modulation schemes (BPSK, 4QAM, 16 QAM) are used. In order to guarantee given bit error rate, every transmission mode (combination of coding rate and modulation) can be used above a specific SNR level. If $BER = P_b$ then γ values for different modes can be calculated using the inverse of (3) and (4). These levels are $\gamma_1 = \frac{1}{2}(Q^{-1}(P_b))^2$ for BPSK, $\gamma_2 = (Q^{-1}(P_b))^2$ for 4QAM and $\gamma_3 = 5(Q^{-1}(\frac{4}{3}P_b))^2$ for 16QAM. Best spectral efficiency can be achieved if BPSK is used in the SNR range $[\gamma_1, \gamma_2)$, 4QAM in $[\gamma_2, \gamma_3)$ and 16QAM is used in $[\gamma_3, \infty)$. If the channel probability density function is $\rho(\gamma)$ than the probability of not sending anything is $P_o = \int_0^{\gamma_1} \rho(\gamma)d\gamma$. Similar equations can be used for calculating the probabilities P_{BPSK}, P_{4QAM} and P_{16QAM} that BPSK, 4QAM or 16QAM are used. We form pairs of obtainable bits per symbol and the probability for those bits per symbol to be obtained. We denote those pairs as $(0, P_o), (1, P_{BPSK}), (2, P_{4QAM}), (4, P_{16QAM})$. These pairs can be viewed as the pdf's of the obtainable bits per symbol. Then for a given outage probability P_{out} the obtainable bits per symbol at that outage probability can be calculated as the minimum value of the bits per symbol R_o for which the cumulative probability of all rates (in bits per symbol) lower than R_o is less than P_{out}. This approach can be extended to several consecutive channel coherence intervals by finding the pdf of the obtainable bits per symbol for the coherence intervals included (convolution can be used in order to obtain this pdf). When coding rates lower than $r_c = 1$ are used, a change in the calculation of γ levels for different modes is needed, which can be done by numerical means in cases where no inverse formulas are available.

The number of obtainable bits per symbol for specific outage probability is dependent on the outage probability and maximal allowed BER. In the first part of the algorithm we propose a way to relate the outage probability and the BER. We assume that the BER is such that the packet error probability for video packet with average length due to outage is equal to the packet error probability when there is no outage. Only packets sent during the coherence intervals when the channel state information is not available to transmitter are included in this calculation. For example, if SQCIF video sequence is used and the frame is transmitted during six coherence intervals, then five video packets will be included in the aforementioned procedure, since for the transmission of the first packet, we assume that the transmitter has channel state information. Next, the number of obtainable bits for the coherence intervals for which the channel gain is unknown (R_{out}) is calculated for the specified BER and outage probability. The number of bits in the current coherence interval is then added to R_{out} (the modes chosen at physical layer, for the first coherence interval, guarantee the same maximal BER). Having the triplets $\{P_{b_i}, P_{out_i}, R_{out_i}\}$, for I chosen values of P_{out} the following optimization process is carried out at the transmitter:

$$\min_{\mu, P_{out_i}, P_{b_i}} D = E[(f_n^j - \tilde{f}_n^j)^2] \tag{5}$$

$$B(\mu) \leq R_{out_i}$$

Here $B(\mu)$ is the number of bits used by the video encoder to encode the video frame. At the end of the opti-

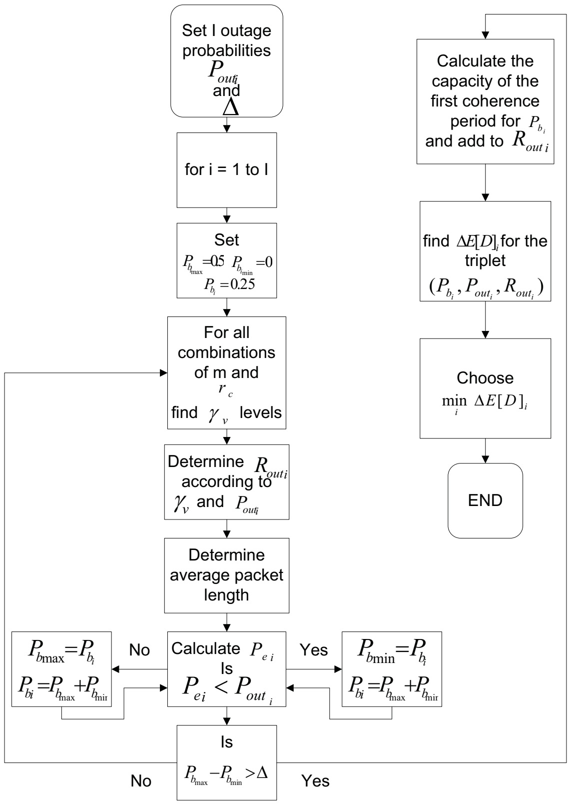

Fig. 1. Flow diagram

mization process coding modes for source coding and triplet $\{P_{b_i}, P_{out_i}, R_{out_i}\}$ that offer minimal expected distortion are chosen. The optimization process in (5) can be done using any optimization approach, for example Lagrangian optimization [3].

The proposed algorithm consists of the following steps:

1) *Set I values for P_{out_i} and set value for Δ (stopping criterion).*
2) *For $i = 1 : I$*
   a) *Set values for $P_{b_{i_{max}}} = 0.5$, $P_{b_{i_{min}}} = 0$, and starting value $P_{b_i} = 0.25$ for the BER.*
   b) *Find thresholds γ_v for all v, that guarantee BER lower than P_{b_i}.*
   c) *Determine the number of obtainable bits R_{out_i}, using γ_v and P_{out_i}.*
   d) *Determine the average packet length L as the ratio between R_{out_i} and the number of packets for which there is no transmitter channel state information. If this value is lower then 100, set the average packet length to 100.*
   e) *Calculate the average packet error probability $P_{e_i} = 1-(1-P_{b_i})^L$ when the system is not in outage. If P_{e_i} is higher then P_{out_i}, set $P_{b_{i_{max}}} = P_{b_i}$, $P_{b_i} = (P_{b_{i_{max}}}+P_{b_{i_{min}}})/2$, else, set $P_{b_{i_{min}}} = P_{b_i}$, $P_{b_i} = (P_{b_{i_{max}}} + P_{b_{i_{min}}})/2$.*
   f) *If $P_{b_{i_{max}}} - P_{b_{i_{min}}} > \Delta$ go to step (b).*
   g) *Calculate the maximal number of bits in the first coherence interval that can be obtained at BER not higher than P_{b_i}. Add this number of bits to R_{out_i}.*
   h) *Given the triplet $\{P_{b_i}, P_{out_i}, R_{out_i}\}$, choose the optimal source coding modes that minimize $E[D]$ (expected distortion) and do not use more bits then R_{out_i}. Save the chosen modes for each i, and set $E[D]_i = E[D]$.*
3) *Choose $\min_i E[D]_i$ and use the saved parameter for that i, to encode the sequence*

The value 100 in step 2) is the average packet length of a video packet when coding a slice at the quantization level offering lowest resolution. The flow diagram of the algorithm is shown in Fig. 1.

III. Simulation results

In all simulations we use the basic version of H.263 [10]. In the coding process the video frame is divided into macroblocks and every macroblock is coded independently. A single macroblock consists of one 16x16 luma part and two 8x8 chroma parts. The packetization is done in error resilient manner, by dividing every video frame into slices. Here a slice is equal to a row of macroblocks. During the coding of the video frame, the encoder can choose between INTRA and INTER coding for every macroblock, and can choose the quantization level for every video slice. The quantization levels are chosen from a set of allowable quantization levels $Q = \{10, 17, 27\}$. Simulations are done using the Foreman SQCIF video sequence composed of 298 video frames. The video sequence has maximal video quality measured in Peak Signal to Noise Ratio (PSNR) of 32.94 dB, obtained by using the quantization level $q = 10$, and assuming that there are no channel errors. At the decoder, simple error concealment is used. In case of error, this error concealment method copies the pixel at the same spatial location in the previous decoded video frame.

In the simulations Rate Compatible Punctured Convolutional Codes (RCPCC) with coding rates $r_c \in \{\frac{4}{5}, \frac{2}{3}, \frac{4}{7}, \frac{1}{2}\}$ is used. The convolutional code has two generator polynomials 171 and 133 and the constrained length of the encoder is $k = 8$. For all simulations hard decoding is used.

In our simulations the modulation types used are from the set $M = \{BPSK, 4QAM, 16QAM, 64QAM\}$.

We model the wireless channel as a block fading Rayleigh channel. The coherence interval of the channel is set to $\frac{1}{6}$ of the duration of one video frame.

In order to evaluate the performance of our algorithm we compare it to algorithm that uses a hypothetical transmitter which knows the channel states for the whole duration of the video frame and has more capabilities at the physical layer, having the option to adjust the power level in different coherence intervals. In this algorithm, which we call JSCCPM [6], the physical layer supplies the application layer with 5 different budgets that correspond to the following bit error rates $\{0.05, 0.01, 0.005, 0.001, 0.0005\}$, and then the video coder finds the source parameters for each value of BER resulting in the lowest expected distortion, $E[D]_i$. The parameters obtained for the lowest $E[D]_i$ are used for transmission.

JSCCPM differs from the newly proposed algorithm in two aspects: its outage probability is equal to zero and it has the ability to use power adaptation. In order to get better insight in the performance of the proposed algorithm, we use yet another algorithm for cross layer optimization of video transmission. This algorithm jointly optimizes the video and physical parameters for every coherence interval. The algorithm maps a slice to a video packet, so constant parts of the video are mapped into strictly defined channel intervals. Therefore, we call this algorithm “algorithm with constant parts”.

In the optimization process for the newly proposed algorithm, we used five different values for the outage probability $P_{out} \in \{0.25, 0.1, 0.05, 0.025, 0.005\}$.

In Fig. 2, a comparison of the performance (measured in terms of PSNR) of the proposed algorithm and the other two algorithms, at a mean channel SNR of 10 dB, is shown.

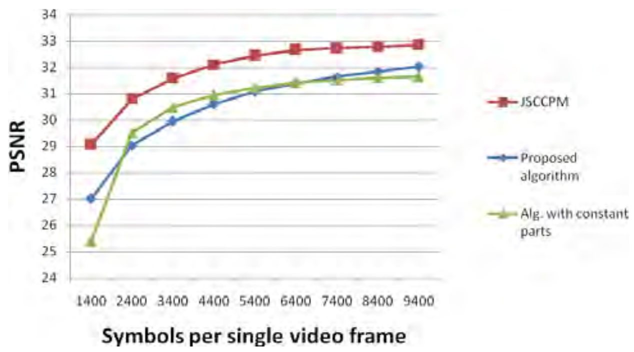

Fig. 2. PSNR vs number of symbols in a single frame

From Fig. 2 it can be seen that the proposed algorithm shows worse performance than JSCCPM. Still, the difference is within 2 dB and reduces to 1 dB for high symbol rates. The proposed algorithm shows better performance compared to the algorithm with constant parts, in the regions of low and high available resources. This is due to the ability of the proposed algorithm to send the video packets for different parts of the video frame in different coherence intervals, so that in the area where extra bits are needed to improve the quality of those video parts improvement in the performance of the received video is achieved. In the remaining region, the use of joint source and channel coding for different coherence intervals, brings improvement for the algorithm with constant parts, compared to the proposed algorithm.

In Fig. 3 we compare the performance of the three algorithms in terms of the mean channel SNR, assuming symbol rate of 2500 symbols per video frame. This figure is for the region of available symbols per video frame where the proposed algorithm shows somewhat worse performance than the algorithm with constant parts. Still, these two algorithms perform closely over the entire region of mean channel SNR. When compared to the JSCCPM algorithm we can see that at high SNR the performace loss is very small. At low SNR the performance loss is significant.

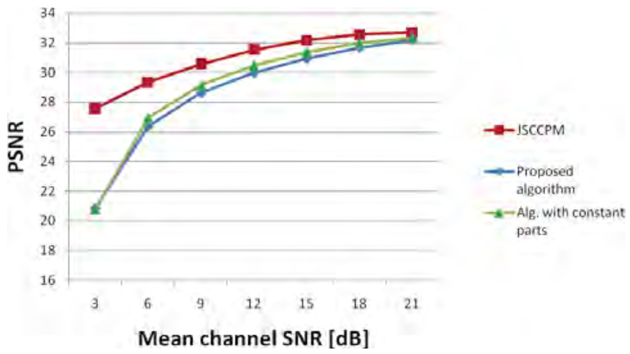

Fig. 3. PSNR vs mean channel SNR

IV. Conclusion

We propose an algorithm for real time video transmission over a wireless channel. In the transmission system causal CSI is available at the transmitter and channel coherence interval is shorter than the transmission interval of a video frame. Encoding procedure is carried out at the beginning of the video frame transmission interval. In the coding process triplets of BER, outage probability and number of bits obtainable at that outage probability are presented to the video encoder. The encoder determines the optimal video coding mode, for each given triplet. The triplet resulting in minimal expected distortion is chosen for coding and transmission. At reasonably high symbol rate and SNR, the proposed algorithm is capable of performing within 1 dB in PSNR of the received video, compared to the algorithm that has ideal knowledge of CSI at all times.

References

[1] A. Argyriou: Distortion-Optimized Video Encoding and Streaming in Multi-Rate Wireless LANs, *IEEE ICASSP*, March 2008, p.p. 2169-2172.
[2] M. Schaar, S. Shankar: Cross-layer wireless multimedia transmission: Challenges, principles and new paradigms, *IEEE Wireless Communications*, August 2005, p.p. 50-58.
[3] F. Zhai, Y. Eisenberg, T. N. Pappas, R. Berry, A. Katsaggelos, ”Joint source-channel coding and power allocation for energy efficient wireless video communications,” *Proc. 41st Allerton Conf. Communication, Control, and Computing*, Octomber 2003.
[4] G. Su, Z. Han, K. Liu, ”Multiuser Cross-Layer Resource Allocation for Video Transmission over Wireless Networks”, IEEE Networks, Vol. 20, pp. 21-27, April 2006.
[5] A. Katsaggelos,Y. Eisenberg, F. Zhai, R. Berry, T. Pappas, ”Advances in Efficient Resource Allocations for Packet-Based Real-Time Video Transmission”, *Proceedings od the IEEE*, Vol. 93, No. 1, January 2005.
[6] S. Pejoski, V. Kafedziski, ”Error resilient wireless video transmission”, Procceding of ETAI 2009 National COnference (in Macedonian), September 2009, Ohrid.
[7] R. Zhang, S. Regunathan, K. Rose, ”Video Coding with Optimal Inter/Intra-Mode Switching for Packet Loss Resilience”, *IEEE Journal on selected area in Communications*, Vol. 18, No. 6, June 2000, pp. 966-976.
[8] Y. Zhang, W. Gao, Y. Lu, Q. Huang, D. Zhao, ”Joint Source-Channel Rate-Distortion Optimization for H.264 Video Coding Over Error-Prone Networks”, *IEEE Transactions on multimedia*, Vol. 9, No. 3, April 2007, pp. 445-454.
[9] J. Hagenauer: Rate-compatible punctured convolutional codes (RCPC codes) and their applications, IEEE Transactions Communications, Vol. 36, April 1988, p.p. 389 - 400.
[10] Available at: http://euler.slu.edu/ fritts/mediabench/mb2/index/html