Video Transmission Over Slowly Fading Channels Using Diversity and Symbol of Power Adaptation

Slavche Pejoski and Venceslav Kafedziski

Abstract—**A novel method for video transmission over fading channels, that minimizes end-to-end video distortion, is presented. The method is based on the use of diversity and optimal allocation of the number of symbols used for transmission of each packet. The impact of channel errors is taken into account using outage probability. Optimization is performed using Lagrangian method. Similar procedure is performed to optimize the power allocated to each packet. Symbol adaptation shows superior results to power adaptation.**

Keywords—**video, cross-layer adaptation, outage capacity, symbol allocation, power allocation.**

I. Introduction

Video transmission is one of the most popular services in the telecommunications world. Much research has been performed in this field, especially for wired channels, but wireless channel still remains in the focus of the research community. The disadvantages of wireless compared to wired transmission are twofold: the available resources are very limited, and the channel conditions change very often. On the other hand, video coding in wireless environment still has its basic features: it is very sensitive to delay and the bit rate needed to encode a single video frame differs significantly from a frame to a frame because it depends on the particular video sequence. This makes the wireless video transmission quite a challenging task, so that optimal use of all available resources and system adaptation to current conditions are necessary. The optimization process is especially needed in real time transmission where the delay should be very low, so retransmission is impossible to be carried out most of the time. In those situations, usually, a combination of forward error correction coding (FEC) and interleaving is employed [1], [2]. In [1] perfect interleaving is employed, which is impossible to do in delay constrained environment. Authors in [2] propose the use of strict set of Reed Solomon (RS) codes that are optimal only for a single SNR value.

The adaptation process has been studied in [1] and [3]. In those works Lagrangian based optimization is used, which is employed as the optimization algorithm in our work, as well. In [3] the channel is assumed to be known and adaptation is done for parts of the video frame, which is not the case here. Authors in [1] use binary modulation only, and the full effect of symbol adaptation is not considered. Review of the achievements in the aforementioned field is presented in [4], [5], [6].

In our previous work [5], we showed that the use of the available channel diversity leads to improvement in the system performance. In that work, we pointed to one of the possible ways to achieve optimal use of the available diversity, based on approximately universal codes. The approach of using approximately universal coding is explained in detail in [7]. In [5] we characterized the slowly varying channel, with no information about the instantaneous channel gain at the transmitter. There, we used the probability density function (pdf) of the channel mutual information to calculate the outage probability of different source coding modes and choose the ones that maximize the end-to-end quality. In the process of choosing video coding modes the transmitter knows that bits obtained from encoding each video slice are mapped to a specific transmission packet. In [5] we assumed that the number of symbols that can be used for transmission of each packet is fixed. But, using the same number of symbols to transmit each packet is suboptimal. The suboptimality is especially pronounced in the region of low symbol rates. Here we use the same approach of calculating the pdf of channel mutual information, mapping parts of the video frame to different video packets and choosing the optimal video coding modes for parts of the video contained in each packet, but extend the work of [5], by including optimal allocation of the available symbols per packet. In order to accomplish this task, we use a nonlinear optimization algorithm based on the Lagrangian optimization.

An optimization algorithm for bit allocation in quantization can be found in [8]. We modify this algorithm to using symbols instead of bits and using end-to-end measure of video quality. This modification requires creation of curves of dependence of the measure of video quality in terms of the number of symbols used for transmission, for each video slice. Then using the calculated curves, Lagrangian optimization is performed. It is known that Lagrangian optimization finds optimal or near optimal solution, which is explicitly shown in [8]. As a measure of the quality of the transmitted video we use the expected end-to-end distortion. The approach is explained in [9] and [10].

We also apply the adaptation to optimal allocation of transmission power to different packets, where we keep the number of symbols per packet fixed, and we compare the two approaches.

The paper consists of four parts. In the second part a description of the transmission system and the formulation of the algorithm itself are given. In the third part, simulation

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

setup and results are shown and in the last part concluding remarks are presented.

II. System Description

The system we use here is the same as the one in [5] and is shown in Fig. 1. At the transmitting side it consists of video encoder, channel encoder, transmitter and controller.

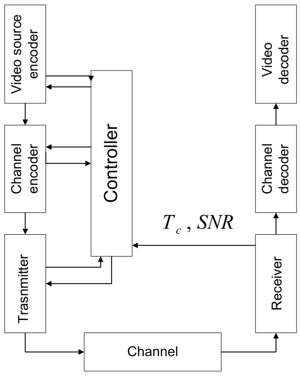

Fig. 1. Video transmission system

In our system the transmitter can have access to certain amount of information about the previously arrived video packets at the receiver. The information delay is dependent on the delay in the transmission system. The transmitter also knows the statistics of the wireless channel i.e. it is familiar with the coherence interval and the probability density function of the channel gain.

The adaptation of system parameters is always based on choosing the optimal modes for different system parameters. In order to evaluate the influence of the different encoding options, an evaluation measure for the received video quality is needed. Here we use the end-to-end expected distortion. The expected distortion is very suitable, because it is calculated at the transmitter and it considers the error probability of video packets and the discrepancy between the available reference frames at the transmitter and the receiver. For calculating the expected distortion of pixel i in video frame n, we use the following expression:

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

where f_n^i is the value of the i-th pixel in the n-th video frame, and $\tilde{f}_n^i$ is the value of the i-th pixel of the n-th video frame at the receiver. To calculate the expected distortion in Eq. (1) recursive algorithms can be used. Two such algorithms are described in [9] and [10].

In the calculation of the expected distortion two types of distortion are considered. The first one comes from the quantization process and the second one comes from the packet loss in the channel. These two parts are combined and are considered together because they are both influenced by the same parameters. For example, the choice of quantization level dictates the quantization error and the number of bits which in turn affects the error probability that influences the distortion from packet losses. The distortion due to transmission depends on the packet error probability that makes this parameter important. The packet error rate in wireless systems is given by [7]:

$$P_e = P_{out}P(error|O) + (1 - P_{out})P(error|O^c) \tag{2}$$

where O is the outage event, O^c is the no outage event, $P(error|O)$ is the error probability when the channel is in outage, and $P(error|O^c)$ is the error probability when the channel is not in outage. Similarly to what we did in [5], we assume that our channel coder is a capacity achieving one and that it uses powerful codes that drive $P(error|O^c)$ to zero. Based on these assumptions, the packet error rate becomes $P_e = P_{out}$. For wireless channel in which the coherence interval is L times shorter than the time necessary for the transmission of a single frame and the whole available diversity is used, the outage probability is calculated according to [11]:

$$P_{out}(R) = P\left(\frac{1}{L}\sum_{i=1}^{L}\log_2(1 + |h_i|^2 SNR) < R\right). \tag{3}$$

Eq. (3) allows us to map the channel conditions to the pdf of the channel mutual information. The pdf of the mutual information can be used to calculate the error probability for different source coding modes.

The novelty of this work comes from adaptation of the number of symbols used by different video packets, that are used for transmission of a single video frame, such that the distortion of the video sequence at the receiver is minimized. The algorithm used to allocate the available resources to different video packets is based on nonlinear optimization process, carried out by Lagrangian optimization. Before explaining the allocation process it is necessary to point out that symbol adaptation over different coherence intervals in a time varying channel, when the diversity is not used, brings very low benefit and great complexity. The utilization of the whole available diversity makes the share of the symbols from different channel intervals equal, resulting in the same expression for the mutual information given by Eq. (3) and hence makes the optimization process of symbol adaptation much simpler.

The optimal use of the available diversity for video transmission is illustrated in Fig. 2, where $\gamma_i = |h_i|^2$. In this figure, time flow starts from the upper left corner and ends at the lower right corner, it is always horizontal, and when it reaches the end of the interval it continues from the left side of the next coherence interval. In the example shown in this figure we assumed that the wireless channel coherence interval is one third of the frame duration. The borders of packets when equal number of symbols are allocated to all video packets are shown in black color. Then Δ_s symbols are reallocated form packet $P2$ to packet $P1$, and the time positions of those symbols are placed in such a manner, that the pdf of channel mutual information per symbol doesn't change. The new borders of packets $P1$ and $P2$ are shown in red.

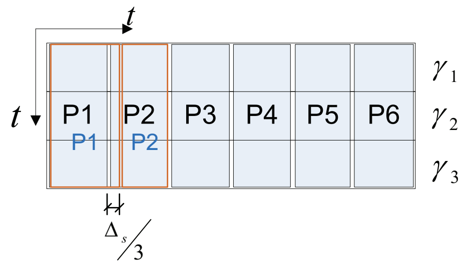

Fig. 2. Optimal use of the available diversity and symbol reallocation

The Lagrangian optimization is a well known method of nonlinear optimization (described, for instance, in [12]). The application of the Lagrangian based algorithm for optimal bit allocation to different subbands in subband coding was first described in [8]. Here, we use this algorithm for optimal allocation of the available symbols to different packets. The algorithm finds an accessible point that lies on the convex hull of the curve that describes the expected distortion in terms of the number of symbols used for transmission, and is closest but lower or equal to the number of available symbols. If the point found during the Lagrangian optimization process has lower number of symbols than the available number of symbols, additional optimization takes place. This additional optimization allocates the rest of the symbols in a greedy manner, i.e. it allocates them to the packet, for which the additional symbols bring largest performance improvement. Obviously, in order for the optimization to take place, curves for expected distortion in terms of the number of used symbols have to be created for each video packet. In the process of curve creation, for every option of used symbols, optimization of the modes used by the video coder must be performed. This additional optimization can be very tedious if a large number of source coding options are available, since it is proportional to the number of source coding options.

We apply the concept of adaptive resource allocation to power allocation as well. In this case, the number of symbols per packet is kept constant, but the available power is allocated to particular transmitted packets in such a way that minimal expected distortion is obtained. For this procedure, we again use Lagrangian optimization. In this case, curves of dependence of the expected distortion in terms of the power used to transmit each packet are created.

III. Simulation Results

In our simulations we use the basic H.263 coder [13]. During the encoding process every video frame is divided into macroblocks, that are independently encoded. Each macroblock consists of 16x16 luma part and two 8x8 chroma parts. In order to obtain better resilience to transmission error the video frame is divided into slices which are independently packetized. Hence, even in situation when a loss of slice occurs, the decoding process for the current video frame can continue. In this work we assume that every slice consists of one row of macroblocks. During the encoding procedure, the encoder can choose between INTRA and INTER modes for every packet, and the quantization level for every slice. The quantization levels are chosen from the set $Q = \{10, 17, 27\}$, which is a subset of the set of all available quantization levels. To get the results we used 298 video frames of the Foreman SQCIF video sequence with maximal receiver video quality measured in Peak Signal to Noise Ratio (PSNR) of 32.94 dB, obtained using quantization level $q = 10$ and assuming no errors during the transmission. For our simulations we used simple error concealment. This concealment method conceals the lost pixels by copying them from the same spatial locations in the previous decoded frame, when packet loss occurs. In all our simulations we anticipate situation with no delay for the information about loss of already sent packets i.e. when coding the current video frame, the video coder knows all the packets that were lost for all video frames in the past. This assumption does not change the conclusions, because the situation with a finite delay can be controlled by the algorithm for calculation of the expected distortion, which results in performance shift only.

In our simulations we use a block fading channel model. The duration of the coherence interval is set to $T_c = 33/6 = 5.5ms$ i.e. there are 6 coherence intervals during the transmission interval of a single video frame. We use the maximum available diversity order 6 if not stated otherwise. The procedure for obtaining the probability density function of the channel mutual information for different diversity order can be found in [5].

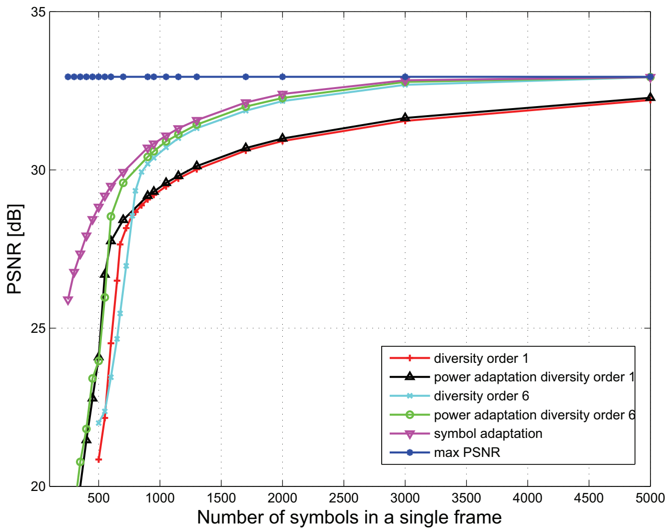

Fig. 3. PSNR vs number of symbols in a single frame

All the simulations are carried out for 20 different channel realizations. The PSNR in all figures is calculated as the mean PSNR over all video frames and channel realizations. In the symbol adaptation algorithm, the symbols are allocated to the packets in portions equal to 10% of the mean available symbols per packet. In the power adaptation algorithm, the power is allocated in portions equal to 10% of the mean power per packet. The resolution of the symbol and power allocation is variable and can be used for trading complexity and performance. Namely, finer resolution implies better performance, but results in larger number of points on the curves and hence in increased complexity.

In Fig. 3 the quality of the received video sequence measured in PSNR in terms of the number of available symbols per video frame is shown. For all simulations shown in this figure the mean channel SNR is set to 16 dB. The figure contains performance curves for both symbol adaptation algorithm and power adaptation algorithm. The simulations for the power adaptation algorithm are performed for diversity order of one and six. The figure also contains curves for quality of the received video for diversity order of one and six, when both the power and available symbols per packet are fixed. It is obvious that symbol adaptation brings the largest benefit. This benefit is most pronounced in the region of low symbol rates. In the remaining region symbol adaptation still has the best achievable performance among the adaptation algorithms shown here, but the difference in performance among different algorithms is quite small. Fig. 3 also shows that symbol adaptation is superior to power adaptation for all available symbol rates. Still, the power adaptation algorithm results in performance gain compared to non symbol/power adaptive algorithms. The superiority of symbol adaptation comes from the high SNR that is used to obtain the results in this figure, namely at high SNR the additional energy doesn't modify the pdf of the channel mutual information as significantly as the additional symbols. Other benefit from using symbol adaptation is that, no matter the allocation, all the available resources are used, which is not the case for power adaptation. Namely, when no power is used for transmission of some video packet i.e. the packet is not sent, the symbols dedicated to that packet are not used.

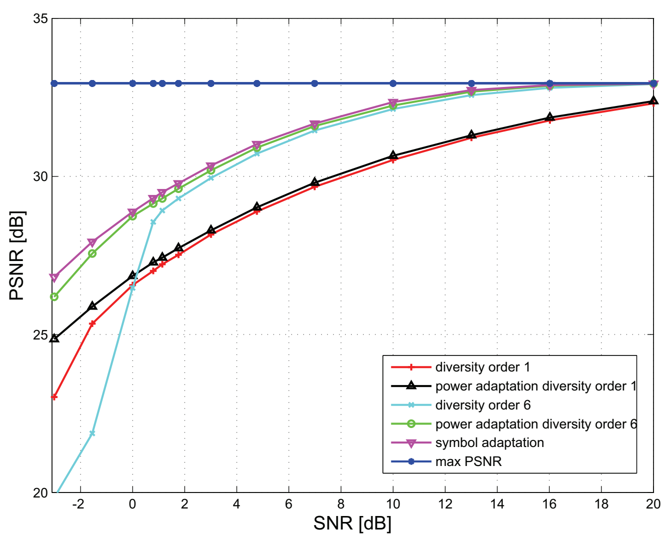

Fig. 4. PSNR vs mean channel SNR

Another comparison of the performance of the proposed algorithms measured in PSNR in terms of channel SNR is made in Fig. 4. In this figure the number of available symbols is kept at 3500 symbols per frame. It can be seen that in this setting the adaptation algorithms have similar performance and they both exceed the performance of algorithms that do not use symbol/power adaptation. Again symbol adaptation brings performance gains compared to power adaptation, but here the difference is not as large as in the constant SNR case.

IV. Conclusion

We present a method for adaptive video transmission in wireless channels. The method uses the pdf of channel mutual information and outage probability to calculate curves of dependence of the video quality on the number of symbols used for the transmission of each packet. For the video part contained in each video packet the optimal mode for source coding is chosen for every point on these curves. Then, symbol allocation procedure based on Lagrangian optimization is carried out. An extension of this concept to power adaptation is also proposed. Comparison between the novel adaptive algorithms and previously developed algorithms with no symbol/power adaptation is performed using simulation. Simulation results show that the adaptive algorithms outperform the constant resource algorithms over the entire region of available resources. The performance gain is highest in the region of limited available resources, where adaptive symbol allocation shows best performance of all.

References

[1] 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.

[2] Q. Qu, Y. Pei, J. Modestino, X. Tian, "Error-resilient wireless video transmission using motion-based unequal error protection and intraframe packet interleaving", *ICIP International Conference on Image Processing*, Vol. 2, pp. 837 - 840, 2006.

[3] A. Argyriou, "Distortion-Optimized Video Encoding and Streaming in Multi-Rate Wireless LANs", *IEEE ICASSP*, March 2008, p.p. 2169-2172.

[4] A. Katsaggelos,Y. Eisenberg, F. Zhai, R. Berry, T. Pappas, "Advances in Efficient Resource Allocations for Packet-Based Real-Time Video Transmission", *Proceedings of the IEEE*, Vol. 93, No. 1, January 2005.

[5] S. Pejoski, V. Kafedziski, "Video transmission on slowly fading channels using diversity", Telfor 2009, November 2009, pp. 307-310.

[6] M. Schaar, S. Shankar, "Cross-layer wireless multimedia transmission: Challenges, principles and new paradigms", *IEEE Wireless Communications*, August 2005, p.p. 50-58.

[7] S. Tavildar, P. Viswanath, "Approximately Universal Codes over Slowly Fading Channels", *IEEE Transactions on Information theory*, Vol. 52, No. 7, pp. 3233-3258, July 2006.

[8] Y. Shoham, A. Gersho, "Efficient Bit Allocation for an Arbitrary Set of Quantizers", *IEEE Transactions on Acoustics, speech and signal procession*, Vol. 36, No. 9, September 1988, p.p. 1445-1452.

[9] 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.

[10] 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.

[11] D. Tse, P. Viswanath, "Fundamentals of wireless communications", Cambridge university press, 2005.

[12] S. Rao, "Engineering Optimization: Theory and Practice, Forth Edition", John Wiley & Sons, 2009.

[13] Available at: http://euler.slu.edu/ fritts/mediabench/mb2/index/html