ICIP'2000, Vancouver, Canada, September, 2000

TRADE-OFF BETWEEN SOURCE AND CHANNEL CODING FOR VIDEO TRANSMISSION

Klaus Stuhlmüller, Niko Färber

University of Erlangen-Nuremberg
Telecommunications Laboratory
Cauerstrasse 7/NT, D-91058 Erlangen, Germany
Email: [stuhl|faerber]@LNT.de

Bernd Girod

Stanford University
Information Systems Laboratory
Stanford, CA
Email: girod@ee.stanford.edu

ABSTRACT

An analytical model of a video transmission system is used to study the impact of its parameters. The optimal trade–off between error resilience using INTRA coding and the channel code rate is determined for given channel parameters by minimizing the expected MSE at the decoder. It is shown that the results strongly depend on the burstiness of the channel.

1. INTRODUCTION

A key problem in the design of a transmission system is the bit allocation between source and channel coding. Often, the underlying transmission system is regarded as a "black box" and the video codec has to cope with whatever bit error rate or packet error rate is offered. This approach is indeed justified if video is added as another application on top of a fixed transmission system. However, current and future transmission systems provide increasing flexibility at the interface to the transport level. For example, the enhanced air interface of the GSM system (EDGE, [1]) will include a flexible link adaptation.

The scope of this paper is to find the optimal trade–off between source and channel coding for a video transmission system. We use a model that is presented in [2, 3] to derive an analytical solution. The optimal trade–off between INTER and INTRA coding has already been derived in [2] and is employed in the results presented here. Similar investigations have been published before for vector quantization [4] and Lempel-Ziv compression [5], but not for motion–compensated video coding.

2. VIDEO TRANSMISSION SYSTEM

As can be seen from Fig. 1, the investigated system consists of three parts, the video encoder, the video decoder, and the *error control channel*, which is defined as the combination of the channel codec and the channel [6].

The video encoder is characterized by its operational distortion–rate (DR) function $D_\mathrm{e}(\beta, R_\mathrm{e})$. I.e., the average distortion D_e is expressed as a function of the average bit rate R_e and INTRA rate β.

After source coding, the compressed video bitstream is prepared for transmission by the channel codec. Often, this involves packetization and some form of *error control*. In this paper we focus on *Forward Error Correction* (FEC). More specifically, we assume an (n,k) Reed-Solomon (RS) block code with a block size of n symbols including $k < n$ information symbols. We use the common choice of 8-bit symbols, i.e., the channel codec operates on bytes. The system parameter we want to focus on is the code rate $r = k/n$. By reducing the code rate, more channel coding redundancy is added to each codeword which improves the error correction capability of the code while reducing the throughput at the same time.

As a channel model we use a two–state Markov model describing errors on the symbol level. As intuitive channel parameters we use the average symbol error rate P_B and the average burst length L_B. Together with the total bit rate R_c, these two parameters completely describe the channel and can be used to, e.g., study the influence of burst errors vs. independent symbol errors. Furthermore, the selected channel model allows to calculate the residual word error rate $P_\mathrm{L}(r, P_\mathrm{B}, L_\mathrm{B})$ after channel decoding from the parameters of the Markov model and the code rate [7]. Thus, the overall performance of the error control channel, including a burst channel and an RS channel codec, can be described analytically.

Finally, the influence of residual errors on the decoded video quality has to be considered. Depending on the error resilience capabilities of the video decoder, a single lost codeword may cause severe image distortion. Fast resynchronization of the bitstream and error concealment are two important issues that can help to mitigate the effect of residual errors. Another important issue is interframe error propagation because errors may be visible over many consecutive frames. The model for interframe error propagation that is used in the following is described in detail in [2, 3].

For the simulation results presented in this paper we use the QCIF test sequence *Foreman* which is H.263 encoded (without options) at 12.5 fps using 125 frames. Each Group Of Block (GOB) is encoded with a header to improve resynchronization. The encoder operates at a constant bit rate R_e which is enforced by a simple rate control on a frame basis. The mode decision is performed according to TMN5. For The total channel bit rate was chosen as $R_\mathrm{c} = 200$ kbps. Unless otherwise noted, the average burst length is set to $L_\mathrm{B} = 8$ while the symbol error rate is selected in the range $P_\mathrm{B} = 0.5 \ldots 10\%$. The block size of the RS code is set to the average GOB size ($n = 222$ bytes). If the RS decoder fails to correct the transmission errors in a block, the video decoder does previous–frame concealment for the affected GOBs.

The averaged Mean–squared–error (MSE) is used as the video quality measure. The overall distortion at the video decoder is the sum of the encoding distortion D_e and the transmission distortion D_v, i.e.,

$$D_\mathrm{d} = D_\mathrm{e} + D_\mathrm{v} \ . \quad (1)$$

Since PSNR is more common to the video coding community, we use $PSNR_\mathrm{d}$ instead of D_d in the plots. The distortion values at the decoder are obtained by averaging over 30 channel realizations.

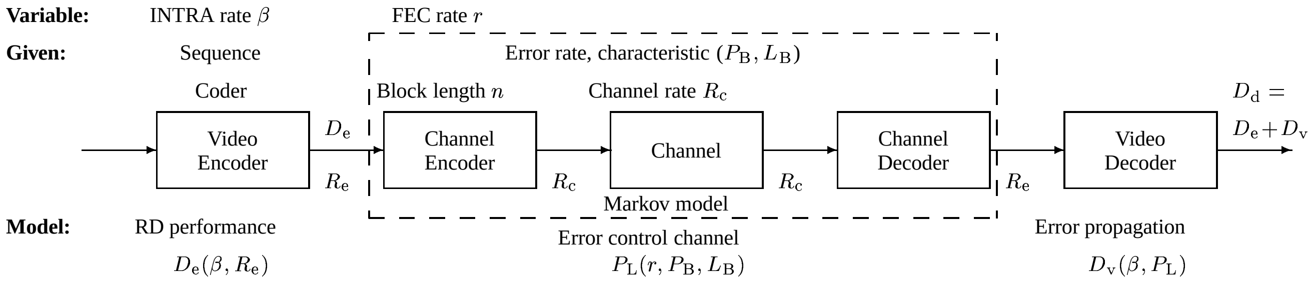

Fig. 1. Video transmission scheme. The overall decoded video quality is denoted D_d.

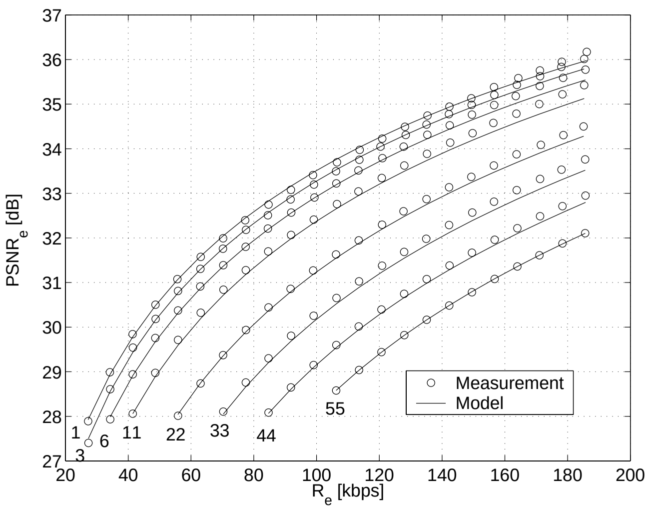

Fig. 2. Distortion–rate curves at the encoder for the test sequence *Foreman* for different percentages of INTRA coded macroblocks β.

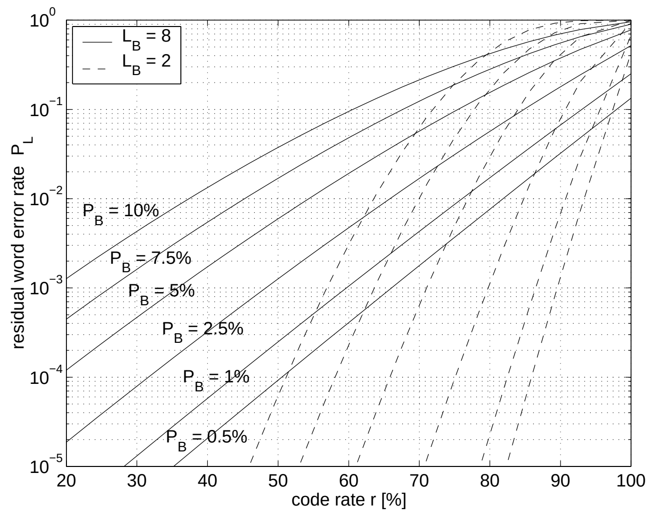

Fig. 3. Residual word error rate P_L for the investigated error control channel ($n = 222$).

3. VIDEO CODEC

The Distortion–Rate (DR) performance of the video encoder is modeled by a simple equation that relates the distortion at the encoder D_e to the relevant parameters. In the considered simulation scenario, there are two parameters with a significant impact on D_e, namely the source rate R_e that is allocated to the video encoder, and the percentage of INTRA coded macroblocks (INTRA rate) β. We have found that the equation

$$D_\mathrm{e} = \frac{\theta}{R_\mathrm{e} - R_0} + D_0 \qquad (2)$$

is suitable to describe the behavior of the video encoder. The variables (θ, R_0, and D_0) are the parameters of the DR model which depend on the encoded sequence as well as on the percentage of INTRA coded macroblocks, β. It can be shown that the relationship with β is approximately linear, i.e.,

$$\begin{aligned} \theta &= \theta_P + \Delta\theta_{IP}\beta \ , \\ R_0 &= R_{0P} + \Delta R_{0IP}\beta \ , \\ D_0 &= D_{0P} + \Delta D_{0IP}\beta \ , \end{aligned} \qquad (3)$$

such that the total number of model parameters is six. According to (3), it is sufficient to measure the DR curves for only two different INTRA rates. Intermediate values can then be obtained by linear interpolation.

Fig. 2 shows that (2) and (3) approximate the DR performance of the video encoder very accurately.

The averaged distortion D_v that is introduced by transmission errors at the decoder has been shown to be [2]

$$D_\mathrm{v} = \sigma_{u_0}^2 \ P_\mathrm{L} \ \left(\frac{\gamma + \beta}{\gamma^2} \ln\left(1 + \frac{\gamma}{\beta}\right) - \frac{1}{\gamma} + \frac{1}{2} \right) \ . \qquad (4)$$

The parameter P_L is the residual word error rate and depends on the channel characteristic as well as on the FEC. The two parameters, γ and $\sigma_{u_0}^2$, have to be estimated for a given video codec, packetization, and video sequence [2]. β is the percentage of macroblocks that are enforced to be coded in Intra mode by the encoder control.

4. ERROR CONTROL CHANNEL

The reliability of a transmission can be improved by Forward Error Correction (FEC), thus lowering the residual word error rate

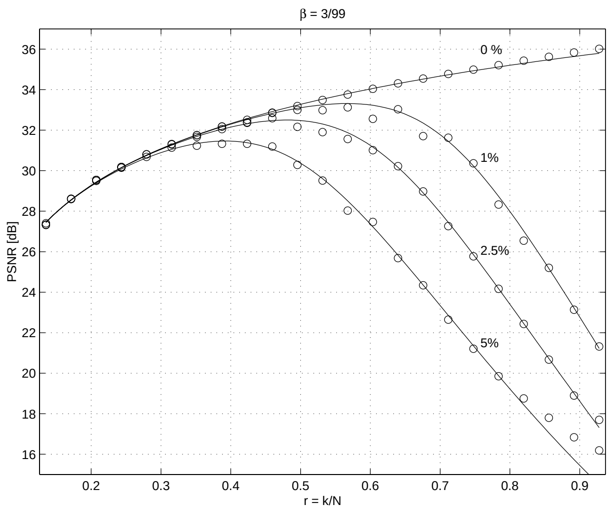

Fig. 4. Measured ('o') and modeled ('—') $PSNR_{\mathrm{d}}$ at the decoder. The channel is characterized by the average burst length $L_{\mathrm{B}} = 8$ and the symbol error rates $P_{\mathrm{B}} = 0\%, 1\%, 2.5\%, 5\%$. The INTRA rate is $\beta = 3\%$.

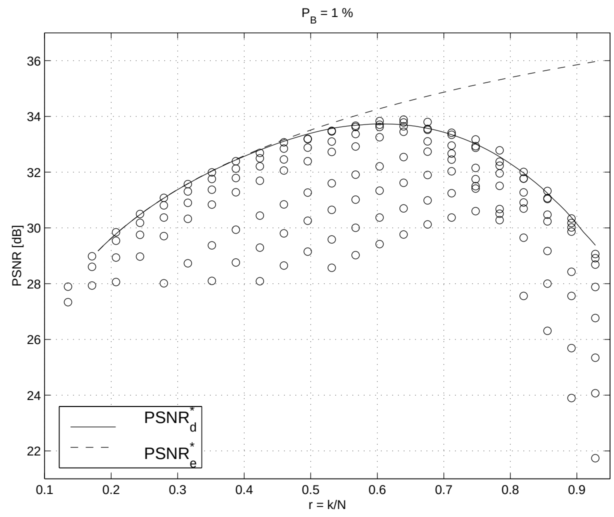

Fig. 5. Optimal $PSNR_{\mathrm{d}}^{*}$ at decoder and corresponding encoder $PSNR_{\mathrm{e}}^{*}$. The INTRA rate β is used as a free parameter for optimization. The channel is characterized by the average burst length $L_{\mathrm{B}} = 8$ and the symbol error rate $P_{\mathrm{B}} = 1\%$.

P_{L} and the additional distortion D_{v} as described by (4). However, in order to maintain a constant channel data rate R_{c}, the available data rate for the source encoder must be reduced to $R_{\mathrm{e}} = rR_{\mathrm{c}}$, where $r \in [0, 1]$ is the channel code rate. This implies that the distortion D_{e} introduced by the source encoder increases. Hence, a trade–off between source coding distortion D_{e} and channel induced distortion D_{v} results. For the optimization of the total distortion D_{d} it is therefore important to understand how much reliability can actually be gained by a certain reduction in code rate.

By using (n, k) RS (Reed–Solomon) codes and *bounded minimum distance decoding* [6] $t_c = \lfloor (n-k)/2 \rfloor$ symbol ($\hat{=}$ byte) errors can be corrected. For other error patterns containing more than t_c symbol errors the RS decoder usually detects that the stream is erroneous and hence error concealment is done in the source decoder.

The *residual word error rate* P_{L} is the probability that a block cannot be corrected. It can be calculated as

$$P_{\mathrm{L}} = \sum_{\kappa=t_c+1}^{n} P_{\mathrm{D}}(n, \kappa) \;, \tag{5}$$

where P_{D} is the *block error density function*. $P_{\mathrm{D}}(n, \kappa)$ denotes the probability of κ symbol errors within a block of n successively transmitted symbols. E.g., for the Binary Symmetric Channel (BSC) with symbol error probability P_{B}, P_{D} is given by the binomial distribution. The P_{D} for a 2–state Markov model which describes a bursty channel can be found, e.g., in [7]. The 2–state Markov model is fully described by the error probability P_{B} and the average burst length L_{B} from which the transition probabilities can be calculated [3]. Note that the model includes the memoryless BSC as a special case by setting $L_{\mathrm{B}} = 1/(1 - P_{\mathrm{B}})$.

The behavior of the investigated error control channel is illustrated in Fig. 3 for a block size of $n = 222$ byte. Note that a variation of the code rate can reduce the residual word error rate P_{L} by several orders of magnitude. This is particular true for small average burst lengths where high reliability can be achieved while maintaining a reasonable throughput.

5. DECODED VIDEO QUALITY

First we study the influence of the channel code rate r on the decoded video quality $PSNR_{\mathrm{d}}$ for a fixed INTRA rate β. Fig. 4 shows that our model approximates the $PSNR_{\mathrm{d}}$ at the video decoder very well. Only for severe channel induced distortion D_{v}, the accuracy of the model is slightly lower. In this case, however, the overall quality is usually far from acceptable anyway such that the achieved accuracy is without practical relevance.

In Fig. 5 the maximum achievable quality at the decoder $PSNR_{\mathrm{d}}^{*}$ is depicted over the code rate r. For each r the INTRA rate β is optimized such that $PSNR_{\mathrm{d}}$ is maximized. The circles mark measurements with different β for the given r. Note that $PSNR_{\mathrm{d}}^{*}$ denotes the upper limit for the given channel, i.e., it is the convex hull of all $PSNR_{\mathrm{d}}$ achievable for the channel with the given code rate r.

Fig. 6 shows the optimal channel code rate r^* for a transmission over burst channels with different average burst lengths $L_{\mathrm{B}} = 2, 4, 8, 16, 32$ and symbol error rates in the range $P_{\mathrm{B}} = 0\%, \ldots, 25\%$. The shown optimal values were found by numerical minimization of D_{d} using the presented system model. Obviously, the optimal r^* is very much dependent on both the symbol error rate P_{B} and the average burst length L_{B}. For increasing error rates the optimal code rate r^* generally decreases corresponding to a stronger FEC.

Interestingly the optimal code rate r^* decreases for bursty channels only to a certain point and increases beyond that point. This is due to the fact that in this case a high INTRA rate β^* is chosen which results in a large R_0 (Eqns. (2) and (3)) thus making it very expensive in terms of D_{e} to further reducing the rate of the video encoder R_{e} by reducing the code rate r.

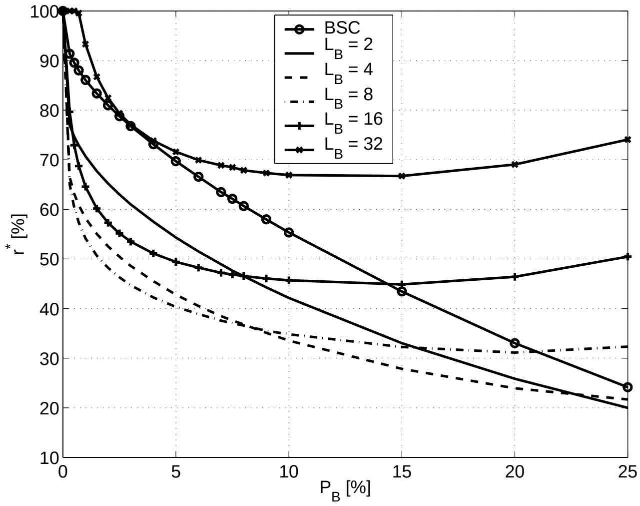

Fig. 6. Optimal code rate r^* for a transmission over channels with different symbol error rates $P_{\rm B}$ and average burst lengths $L_{\rm B}$. The optimal parameters for a channel with independent errors (BSC) are also shown.

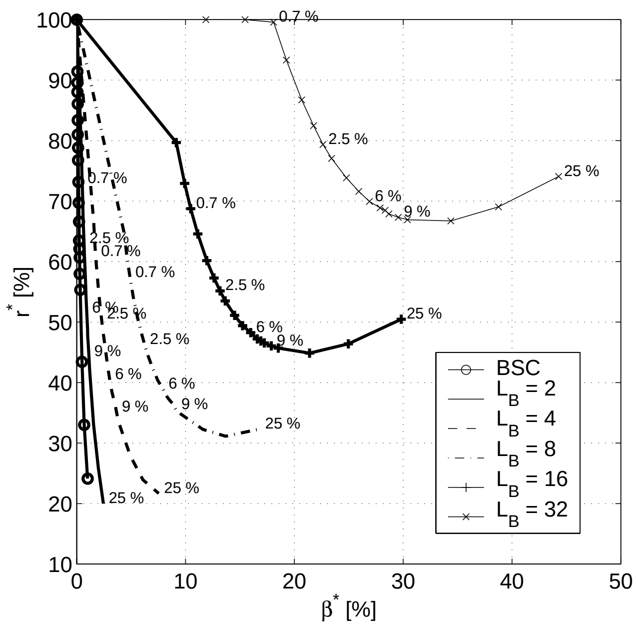

Fig. 7. Optimal parameters β^* and r^* dependent on the symbol error rate $P_{\rm B}$ of the channel for various average burst lengths $L_{\rm B}$ and for the memoryless channel (BSC).

In Fig. 7 the optimal INTRA rate β^* is plotted vs. the optimal code rate r^* in the $\beta^* - r^*$-plane. Both parameters are optimized using the presented system model. Each trajectory corresponds to a given burst length $L_{\rm B}$ while the symbol error rate $P_{\rm B}$ is used as the free parameter that is varied along the trajectory. It can be seen clearly that only for very bursty channels a high INTRA rate β is needed. Note that for a large average burst length $L_{\rm B}$ the optimal code rate is $r^* = 100\%$ for low symbol error rates $P_{\rm B} \leq 0.7\%$. In this case errors occur very infrequently, but if one occurs it is followed by a whole burst of errors. A lot of FEC would be needed to correct this error burst, thus lowering the available rate for the video bitstream in many blocks which are not affected by channels errors at all. This becomes too expensive at some point and it is better to increase the INTRA rate β. However, for non–bursty channels (BSC) the error resilience of the video bit–stream (INTRA) plays a minor role and FEC is very powerful.

6. CONCLUSIONS

We have studied the trade–off between source and channel coding for video transmission. An analytical model was used to show the impact of the channel code rate on the decoded video quality. The transmission over bursty channels has been compared to the memoryless case. The results show that for a memoryless channel FEC is very important whereas error resilience, i.e. INTRA coding, is not so important in this case. In contrast, for bursty channels the use of FEC is limited and the INTRA update is essential.

7. REFERENCES

[1] A. Furaskär, S. Mazur, F. Müller, and H. Olofsson, “EDGE: Enhanced data rates for GSM and TDMA/136 evolution,” *IEEE Personal Communications*, vol. 6, no. 3, pp. 56–66, June 1999.

[2] N. Färber, K. W. Stuhlmüller, and B. Girod, “Analysis of error propagation in hybrid video coding with application to error resilience,” in *Proc. IEEE Int. Conf. on Image Processing (ICIP)*, Kobe, Japan, Oct. 1999, pp. 550–554.

[3] K. W. Stuhlmüller, N. Färber, M. Link, and B. Girod, “Analysis of video transmission over lossy channels,” *IEEE Journal on Sel. Areas in Com.*, vol. 18, no. 6, pp. 1012–1032, June 2000, Special Issue on Error Resilient Image and Video Transmission.

[4] B. Hochwald and K. Zeger, “Tradeoff between source and channel coding,” *IEEE Trans. on Inform. Theory*, vol. 43, no. 5, pp. 1412–1424, Sept. 1997.

[5] G. Buch, F. Burkert, J. Hagenauer, and B. Kukla, “To compress or not to compress?,” in *Proc. 5th CMTC in conj. w. GLOBECOM’96*, London, UK, Nov. 1996, pp. 198–203.

[6] S. B. Wicker, *Error Control Systems*, Prentice-Hall, Englewood Cliffs, NJ, 1995.

[7] E. O. Elliott, “A model of the switched telephone network for data communications,” *Bell. Syst. Techn. J.*, vol. 44, pp. 89–109, Jan. 1965.