BIT ENERGY DISTRIBUTION EFFECTS IN PROGRESSIVELY ENCODED VIDEO FRAME TRANSMISSIONS

Martin Sehlstedt and James P. LeBlanc

Luleå University of Technology
Division of Signal Processing,
SE-971 87 Luleå, SWEDEN
e-mail: {martins, leblanc}@sm.luth.se

ABSTRACT

We investigate physical layer opportunities resulting from multimedia streams which use progressive coding. With such progressively encoded video frames the location of the first bit error is more important than the overall bit error probability. This makes the expected error free length a reasonable metric for performance optimization. One physical layer opportunity is the adjustment of bit energy within data frames. For wireless links nonuniform bit-energy distributions result in a flexible solution that may be used in conjunction with channel coding. We identify system gains to be achieved with nonuniform bit energy distribution optimizations and also show that a simplified frame truncation technique may be used without significant loss of quality.

1. INTRODUCTION

High bit-rate connections makes multimedia over the Internet possible. The video stream consumes the most resources so we concentrate on this within a system containing both wired and wireless links as in Figure 1.

Figure 1: Streaming video over network with wireless link.

For wired networks the SNR is high and the main problem is congestion. A wireless link however introduces into the network new sources for channel unreliability, such as fading and interference. We consider how physical layer processing could be used to achieve graceful degradation of the video quality as channel conditions vary.

Starting with the source, to make graceful degradation possible the source video must be available at different quality levels. There are a variety of ways to do this is e.g. *par-*

Figure 2: Perceptual importance vs. position for progressive data.

allel encoders, *pyramidal encoder* [1], *adaptive code rate* [2], or *progressive encoder* [3].

In the case of *parallel encoder* the source provides different qualities in different streams and the user connects to the stream with the desired quality. A *pyramidal encoder* also consists of several streams but they are organized as one basic stream and several refinement streams so the user connects to the basic stream and a number of refinement streams to get the desired quality. The *adaptive code rate* uses a feedback from the decoder to adjust the code rate and channel coding to achieve best performance at all times. *Progressive encoding* is a recent image compression technique where data for an image are ordered in perceptual importance, see Figure 2. For our purposes here, we consider video to be a succession of such progressively coded images.

All the different methods create different types of problems in network transportation. The parallel encoder generates a high load on network since several quality levels may have to be present at the same time. This problem is somewhat reduced with the pyramidal encoder, but the transportation should prioritize the basic layer. Two possible protocols for this are Differentiated Services, (DiffServ) [4] and Selective Truncation Internetnetwork Protocol, (STRIP) [5]. With DiffServ different quality levels are sent in separate IP-packets and in the case of congestion a router is allowed to drop the lower priority packages. However, with this method packet reordering can cause problems in applications requiring low latency, such as video conferencing, if image frame packets are reordered by the network the decoder must buffer data to restore the order, introducing

Partially supported by the Swedish IT Institute Internet3 Program, and EU Mål 1.

delay. The packet reordering problem is smaller in STRIP, since data for one frame remains in one packet but the data is ordered within the packet with high priority basic data first followed by lower priority enhancement layers. In the case of congestion the STRIP aware routers are allowed to remove the low priority data part of the packet. For the adaptive code rate the major drawback is the delay in adaption due to the delay of the feedback.

With *progressive encoded* data the encoder transmits at full rate and if the channel can not support full resolution transmission, due to congestion or radio link affects, transmission quantization can be done by truncation of frame data in the progressive stream. This requires that the packets with progressive data are “tagged” to mark them as truncatable. One such proposed method to provide such functionality is known as Hints and Notifications [6]. The drawback with progressive video coding is the use of variable code word length in the source compression which means that data decoded after the first bit error becomes essentially useless. The error free decoded data length decides the quality. The result is that the location of the first error is more important than the overall bit error probability. One channel coder that addresses this problem is FEED [7]. It focuses on determining the location of the first non-correctable error. The advantage with progressively coded data in the case of graceful degradation is that it allows for a more continuous scale of quality since the decoded quality depends on the decoded length of the frame data.

Herein we investigate how the radio interface in a wireless link could identify the tagged progressive data and use nonuniform bit energy distributions to achieve unequal error protection to increase the error free length. Possible ways to achieve nonuniform bit energy distributions are to use bit loading in OFDM [8], HQAM [9] or the punctured convolution codes in FEED. Our analysis herein ignores any specific channel coding per sé, but the ideas presented may be applied in concert with such channel coding.

2. MODEL

We assume that the video coder produces a fixed frame rate of progressively coded video data (e.g. SPIHT [3] video coder) where each frame contains N symbols from a QPSK constellation. In our case $N = 19200$ this corresponds to an image of 320 x 240 pixels compressed to 1/2 bit per pixel.

The radio channel model is assumed to be a flat fading channel with slow fading. This can be modeled with AWGN with slowly varying noise level, which is assumed to be constant during a frame. The wireless link is also assumed to include a low data rate feedback path or some other mechanism to estimate link performance. This simple model allows analytical calculations, reducing the need for simulations.

3. NONUNIFORM BIT-ENERGY AND ERROR FREE LENGTH

With nonuniform bit-energy distributions different symbols necessarily have different error probabilities. We denote the energy in symbol n of the frame as $E_s(n)$ for $n = 0, \dots, N - 1$. We then have

$$E_s(n) \geq 0, \quad n = 0, \dots, N - 1$$

$$\sum_{n=0}^{N-1} E_s(n) = N,$$

which ensures a constant average power using the same energy for each frame. The symbol error probability for each symbol, p_n , is

$$p_n \approx 2Q\left(\sqrt{\frac{E_s(n)}{N_0}}\right) \text{ where } Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-t^2/2} dt.$$

We define the error free length, l , to be the position of the first error. Assuming independent symbol errors

$$Pr(l = n) = p_n \prod_{m=0}^{n-1} (1 - p_m). \quad (1)$$

The expected error free length, L , is given by

$$L = \mathcal{E}\{l\} = \sum_{n=0}^{N-1} n Pr(l = n) + N \prod_{m=0}^{N-1} (1 - p_m). \quad (2)$$

In agreement with intuition it can immediately be seen that $E_s(n)$ is a non-increasing function of n . To see this, consider a strictly decreasing symbol energy $E_s(n)$. It can be shown that swapping any two contiguous elements lowers the expected error free length.

4. OPTIMAL NONUNIFORM BIT-ENERGY DISTRIBUTION

We seek the optimal bit-energy distribution to maximize the expected error free length, L , for a given channel's SNR. The N -vector optimization problem takes place in a subset of $\mathcal{R}^N$, where a closed form of the solution is difficult. However, a numerical *steepest ascent* [10] may be used. In actual operation the optimization can be done off line and stored in a lookup table.

The steepest ascent method is an iterative method originally used for unconstrained optimization, the step update is

$$E_s^{(k)} = E_s^{(k-1)} + \mu \nabla_L(E_s^{(k-1)})$$

where μ is a small constant controlling the step size and $\nabla_L(E_s)$ is the gradient of L

$$\nabla_L(E_s) = \left[\frac{\partial L}{\partial E_s(0)}, \dots, \frac{\partial L}{\partial E_s(N-1)} \right]^T.$$

Figure 3: Optimization in $\mathcal{R}^3$ with projection of gradient to unit energy plane.

It is obvious that increasing the total energy will increase L , so ∇_L , will always point in a direction that needs more total energy. To consider a fixed total energy one must constrain the update to the fixed energy hyper plane. This can be done by partitioning the gradient

$$\nabla_L(E_s) = \nabla_L^P(E_s) + \nabla_L^\perp(E_s)$$

such that $\nabla_L^P(E_s)$ lies in the constant total energy plane and $\nabla_L^\perp(E_s)$ is perpendicular to the same plane. Figure 3 illustrates this partition for optimization in $\mathcal{R}^3$. With this partitioning the constrained update is

$$E_s^{(k)} = E_s^{(n-1)} + \mu \nabla_L^P(E_s^{(n-1)})$$

To calculate the gradient one needs the partial derivatives of the expected error free length

$$\frac{\partial L}{\partial E_s(n)} = \frac{\partial L}{\partial p_m} \frac{\partial p_m}{\partial E_s(m)}$$

where

$$\frac{\partial L}{\partial p_m} = m\Phi_{m-1} - \sum_{n=m+1}^{N-1} np_n \frac{\Phi_{n-1}}{1-p_m} - N \frac{\Phi_{N-1}}{1-p_m}$$

and

$$\Phi_n = \prod_{i=0}^n (1-p_i)$$

and

$$\frac{\partial p_m}{\partial E_s(m)} = \frac{-e^{\frac{E_s(m)}{N_0}}}{2\sqrt{\pi E_s(m) N_0}}.$$

Further, it is possible to show that $\frac{\partial L}{\partial p_m} < 0$ for all $m = 0, \dots, N-1$, so we have $\frac{\partial L}{\partial E_s(m)} > 0$.

4.1. Calculations

The modified steepest descent method was used to find the optimum for a frame with $N = 19200$ symbols for different values of SNR the result is shown in Figure 4. Note that the

Figure 4: $E_s(n)$ as a function of n for some different channel SNR.

Figure 5: Probability density function for first error position in frame for different values of channel SNR using the optimal $E_s(n)$.

bit energy distribution is a decreasing function and that for very low SNR all the energy is used for the first bits, like some sort of “soft” frame truncation. With values for $E_s(n)$ the expected error free length L can be calculated directly using (2) and (1). The probability density function, from (1), for the first error position is shown in Figure 5. We see from these density functions that the expected length will be close to the “truncation” limit in all cases, meaning that the variance in the error free length will be small, reducing flicker in video stream at the receiver.

5. NONUNIFORM BIT-ENERGY DISTRIBUTION VERSUS FRAME TRUNCATION

Seeing that with the optimized $E_s(n)$ the energy tended to be used for the first bits we now consider a “hard” frame truncation. This requires only a single variable optimization, the truncation point, using for example *golden-section search* [11] over the complete frame length. The result of one such optimization is shown in Figure 6 which should be compared to the optimal $E_s(n)$ in Figure 4.

A comparison of the performance of optimized bit energy distribution and optimized truncation is shown in Figure 7. The difference in expected error free length in performance is small, on the order of 0.2 dB in favor of the optimized bit energy distribution. The probability density function for the truncated frames, Figure 8, shows that in this case the variance will also be very small, implying that the truncation may be used instead of the optimal $E_s(n)$ without significant loss of video quality.

Figure 6: Optimal $E_s(n)$ and truncated frame as a function of n .

Figure 7: Performance of optimal $E_s(n)$ and truncated frame as a function of channel SNR.

6. CONCLUSIONS

Considering networks with wireless links for video transmission, the use of progressive video coders would allow for a general approach to graceful degradation and robustness. However, such methods must be used in conjunction with network protocols that are able to denote data as truncatable.

We suggest the use of nonuniform bit-energy distributions in the radio interface which are optimized to improve the expected error free length. As those optimizations require knowledge of the radio link SNR, some SNR estimate must be available.

The optimized $E_s(n)$ yields a bit energy distribution which goes to zero before the end of the frame in lower SNR cases (a "soft frame truncation"). Alternatively, a "hard" frame truncation, in which all the available energy is used evenly for contiguous bits from the beginning of the frame was also investigated. In both cases, the performance was similar, agreeing to within a few percent. Given that, the second order statistics for error free length also closely agreed, the simple frame truncation may be preferable in light of simplified transmitter complexity.

7. REFERENCES

- [1] M-T. Sun and A. R. Reibman, Eds., *Compressed Video over Networks*, chapter 7 Layered Coding, Marcel Dekker, 2001.
- [2] M. P. C. Fossorier, Z. Xiong, and K. Zeger, "Progressive Source Coding for a Power Constrained Gaus-

Figure 8: Probability density function for first error position in frame for different values of channel SNR using the optimal truncation.

sian Channel," *IEEE Trans. on Commun.*, vol. 49, pp. 1301–1306, August 2001.

- [3] A. Said and W. A. Perlman, "A new, fast and efficient image codec based on set partitioning in hierarchical trees," *IEEE Trans. on Circuits and Systems for Video Techn.*, pp. 243–250, June 1996.
- [4] M. Carlsson, W. Weiss, S. Blake, Z. Wang, D. Black, and E. Davies, "An architecture for differentiated services," *RFC 2475, IETF Network Working Group*, December 1998.
- [5] M. Engan, P. Fransson, and O. Schelén, "Selective truncation internetwork protocol: experiments with explicit framing," *Proceedings Ninth International Conference on Computer Communications and Networks*, pp. 220–224, 2000.
- [6] L.-Å. Larzon, U. Bodin, and O. Schelén, "Hints and notifications," in *Conference proceedings WCNC'02*, 2002.
- [7] J. Hagenauer, T. Stockhammer, C. Weiss, and A. Donner, "Progressive source coding combined with regressive channel coding on varying channels," *ITG-Fachbericht*, pp. 123–130, 2000.
- [8] J. A. C. Bingham, "Multicarrier modulation for data transmission: An idea whose time has come," *IEEE Commun. Mag.*, vol. 28, pp. 5–14, May 1990.
- [9] S. Mirabbasi and K. Martin, "Hierarchical QAM: A spectrally efficient DC-free modulation scheme," *IEEE Commun. Mag.*, pp. 140–146, November 2000.
- [10] S. G. Nash and A. Sofer, *Linear and Nonlinear Programming*, McGraw Hill, 1996.
- [11] M. T. Heath, *Scientific Computing: An Introductory survey*, McGraw Hill, 1997.