

ENERGY EFFICIENT WIRELESS TRANSMISSION OF MPEG-4 FINE GRANULAR SCALABLE VIDEO

Cristina E. Costa¹, Yiftach Eisenberg², Fan Zhai², Aggelos K. Katsaggelos²

¹*DIT, University of Trento - Via Sommarive 14, I-38050 Trento (ITALY) - cristina.costa@ing.unitn.it*

²*Northwestern University, Department of ECE, Evanston, IL 60208, USA - {yeisenbe, fzhai, aggk}@ece.northwestern.edu*

Abstract - Fine granular scalability is a coding tool, recently introduced in the emerging MPEG-4 standard, which enables the creation of very flexible scalable video bitstreams. This paper investigates the transmission of fine granular scalable (FGS) video over wireless links, using power management for unequal error protection of the bitstream. In wireless systems, energy may be a limited resource, and a wise use of it is important for system efficiency. An algorithm is proposed which is able to optimally distribute the total available power for the transmission of the enhancement layer, given a distortion or energy constraint. Experimental results demonstrate the performance advantage of the proposed algorithm over fixed power schemes and heuristic approaches.

1 INTRODUCTION

Emerging wireless communication systems have evolved from a voice based service to a multimedia oriented one. This transition has increased the demands and complexity required from transmitting devices. Wireless networks have unique characteristics that distinguish them from more traditional wired networks. These characteristics must be taken into account when designing multimedia systems and, when possible, exploited as an advantage. One of the major obstacles in wireless links is the effects of channel losses on the transmitted data. This, combined with the variability of the channel characteristics and the limitation of resources, such as power, motivates the need for further research into different approaches for data transmission.

When speaking about video coding and transmission, different approaches have been proposed to cope with these limitations. Among them, the use of scalability has been seen as a useful tool and new methods for achieving it have been included in recent standards [1].

Scalability is a coding method that produces an encoded sequence easily capable of accommodating different bitrates. In order to generate a scalable bitstream, the fine granular scalability (FGS) encoder produces two bitstreams, commonly named *Base* and *Enhancement Layers*. The Base Layer (BL) can be decoded independently from the Enhancement Layer (EL), and produces a low quality reconstruction of the video sequence. A higher quality reconstruction can then be achieved by decoding both the Base and Enhancement Layers together.

Currently, *Part 2* of the MPEG-4 standard includes FGS as an encoding tool [1],[2]. The FGS approach differs from traditional layered methods for video scalability because of its capability to achieve a smooth transition between different bit rates. In MPEG-4 FGS, the Base Layer behaves as a standard baseline MPEG-4 compressed bitstream. The difference between the reconstructed BL and the original video sequence is encoded in the Enhancement Layer. Progressive encoding of the EL is achieved by a bit-plane coding of the DCT of the residual image. The DCT transform is performed on a block basis, as is done for the Base Layer, but the data in the EL is

entropy coded one bit-plane (BP) at time. The encoded data is then sent starting from the most significant bit-plane (MSBP) to the least significant bit-plane (LSBP)

Due to its structure, the enhancement layer can be truncated at any point. In this way, rate control can be performed by simply truncating the bitstream in order to adapt it to varying channel characteristics. Due to its inherent scalability and flexibility, FGS also allows complexity scalability and easy resource adaptation depending on the capabilities of video devices. Thus, in addition to video conferencing, FGS is also suitable for video multicast. An interesting overview of applications enabled by FGS technology is given in [3].

With regard to the related work, unequal error protection (UEP) between BL and EL using different techniques, such as ARQ (Automatic Repeat reQuest) and FEC (Forward Error Correction), have been studied in the past [4]. The application of UEP within the EL FGS bitstream was first considered by Schaar et al. in [5], where the Fine-Grained Loss Protection (FGLP) framework was introduced. Based on it, Yang et al. proposed in [6] a “degressive” protection algorithm (DEP) based on FEC for optimally assigning protection redundancy among bit-planes. In [7], Wang et al. studied the problem of rate-distortion optimized UEP for Progressive FGS (PFGS) over wireless channels using prioritized forward error correction (FEC) for the BL and EL. A similar problem was studied in [8] in which the objective was to minimize the processing power for PFGS video given bandwidth and distortion constraints. In [9], a joint FEC and transmission power allocation scheme for layered video transmission over a multiple user CDMA networks was proposed. In that work, scalability was achieved using 3D-SPIHT (wavelet based coding). The objective was to minimize the end-to-end distortion through optimal bit allocation among source layers and power allocation among different CDMA channels. The authors in [10] considered jointly adapting the source bit rate and the transmission power in order to maximizing the performance of a CDMA system subject to a constraint on the equivalent bandwidth. In that work, an H.263+ codec was used to generate the layered bit stream.

In this work, we consider the transmission of an MPEG-4 FGS video sequence over a single wireless channel for a single user. We propose an algorithm that allows UEP of the enhancement layer packets through optimal transmission power allocation. Given transmission power and bit rate constraints, our goal is to minimize the end-to-end distortion by optimally allocating transmission power to the different packets in the EL.

2 FGS IN WIRELESS VIDEO TRANSMISSION

When transmitting in an error prone environment, error protection techniques become of fundamental importance. In

scalable coding, the compressed video is coded into separate bitstreams, which can receive different levels of protection. Typically, the Base Layer is heavily protected so that it is received nearly error free, while a different and lighter protection scheme is used for the Enhancement Layer.

Due to the structure of the FGS EL bitstream, it is also possible to perform unequal error protection within the EL, since the importance of the data within the layer decreases going from the MSBP to the LSBP.

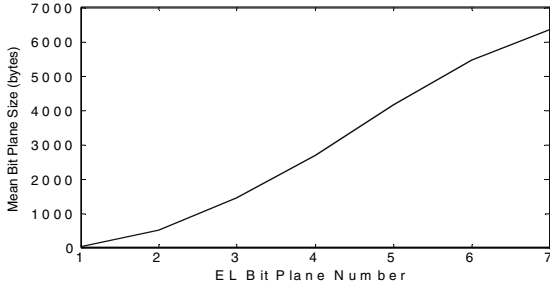


Fig.1 – Average size of the bit-planes for the *Foreman* sequence (QCIF).

By analyzing the EL bitstream, we can note that typically the most significant bit-plane is the smallest in size (i.e., has the smallest number of bits), and bit-plane size increases from the most significant to the least significant bit-plane. As an example, in Fig.1 the average size of the different bit-planes is plotted for the *Foreman* sequence (the BL was coded with bit rate set at 14Kbps). This behavior is due to several factors. It is likely that the first bit plane contains mostly the information related to those macroblocks that have more error in the base layer, such as those with a lot of motion information. This means that the first bit plane contains a greater number of zeros in regions where there are static macroblocks, such as the background or continuous tone areas. Fewer bits are required to encode large regions of zeros. Another reason is due to the fact that the data in the least significant bit planes is less correlated, and thus more difficult to compress with entropy coding.

As a consequence, not all the data in the Enhancement Layer has the same importance. First, it is not possible to decode the data of a bit plane without decoding the preceding bit planes. Secondly, the data of the MSBP also carries more information, with respect to the LSBP. It is possible then to exploit this structure for implementing unequal error protection (UEP) of the enhancement layer data as in [5].

Another important issue in mobile environments is power management. Energy is a precious resource, since batteries have a limited life, and using it in a wise way can greatly improve the overall quality of the transmission. Since the probability of packet lost is directly related to the power assigned to the transmitted packet, we can think of controlling the probability of loss by varying the transmission power in order to achieve unequal error protection of the packets [11].

In this paper, an algorithm is proposed that allows unequal error protection of the enhancement layer packets through optimal transmission power allocation.

3 POWER BASED UNEQUAL ERROR PROTECTION

Taking advantage of the structure of FGS coding, it is possible to implement a prioritization of the enhancement layer packets: a higher level of protection can be given to the

more significant bit-planes, making the protection smaller as the bit-planes become less significant.

Assuming that the BL is always received correctly and that a maximum amount of energy E_{Etot} for the frame is given, our goal is to determine how to allocate the available power in such a way that minimizes the overall distortion.

If we subdivide the EL bitstream into L packets and transmit these packets over the channel, then the total energy used for transmitting the EL is given by the following formula:

$$E_{Etot} = \sum_{l=1}^L \frac{B_E^l \cdot P_E^l}{R}, \quad (1)$$

where B_E^l is the number of bits in the l -th packet, P_E^l is the power used for transmitting it, and R is the channel bit-rate. The minimization problem can be expressed as in the follows:

$$\min_{P_E^l} \{D_B - E[\Delta_E]\}, \text{ s.t. } E_{Etot} = \sum_{l=1}^L \frac{B_E^l \cdot P_E^l}{R}, \quad (2)$$

where D_B is the distortion of the BL and $E[\Delta_E]$ is the expected value of the distortion improvement introduced by jointly decoding the BL and all the correctly received EL packets. If we assume that each packet is successfully decoded only if it and all the previous EL packets are correctly received, we can write $E[\Delta_E]$ as:

$$E[\Delta_E] = \sum_{l=1}^L \prod_{i=1}^l (1 - \rho_E^i) \cdot \Delta_l, \quad (3)$$

where ρ_E^l is the loss probability for the l -th packet, and Δ_l is the distortion improvement introduced by it. By introducing a Lagrange multiplier λ , the solution of the constrained minimization problem can be found by solving the following unconstrained problem:

$$\min_{P_E^l} J = \min_{P_E^l} \left\{ D_B - \sum_{l=1}^L \prod_{i=1}^l (1 - \rho_E^i) \cdot \Delta_l + \lambda \left(\sum_{l=1}^L \frac{B_E^l \cdot P_E^l}{R} \right) \right\}. \quad (4)$$

The average transmission power used by a modulation scheme directly affects the probability of packet loss. We assume that the relationship between the probability of loss for the l -th packet ρ_E^l and the power P_E^l used for transmitting it is known at the transmitter:

$$\rho_E^l = g(P_E^l). \quad (5)$$

The function g can be defined using an analytical model of the wireless channel or determined from empirical measurements. In Section 4, we use an analytical model based on information theoretic results.

The first derivative of the cost function J with respect to P_E^L can be written as:

$$\frac{\partial J}{\partial P_E^L} = \frac{\partial \rho_E^L}{\partial P_E^L} \cdot \prod_{i=1}^{L-1} (1 - \rho_E^i) \cdot \Delta_L + \lambda \frac{B_E^L}{R}. \quad (6)$$

Note that at the optimal solution the derivative in (6) must equal zero. Rearranging equation (6) we can write the following expressions, for $\rho_E^i \neq 1$:

$$\prod_{i=1}^{L-1} (1 - \rho_E^i) = \lambda \cdot f(P_E^L, B_E^L, \Delta_L), \quad (7)$$

from which we can easily obtain, for $j=1, \dots, L-2$:

$$\prod_{i=1}^j (1 - \rho_E^i) = \lambda \cdot f(P_E^L, B_E^L, \Delta_L) \cdot \prod_{h=j+1}^{L-1} (1 - \rho_E^h)^{-1}, \quad (8)$$

where

$$f(P_E^L, B_E^L, \Delta_L) = -\frac{B_E^L}{R \cdot \Delta_L} \cdot \left(\frac{\partial \rho_E^L}{\partial P_E^L} \right)^{-1}.$$

The first derivative of the cost function J with respect to P_E^j can be written, for $\rho_E^j \neq 1$, as:

$$\frac{\partial J}{\partial P_E^j} = \frac{\partial \rho_E^j}{\partial P_E^j} \cdot (1 - \rho_E^j)^{-1} \cdot \sum_{l=j}^L \left(\prod_{i=1}^l (1 - \rho_E^i) \right) \cdot \Delta_l + \lambda \frac{B_E^j}{R} = 0. \quad (9)$$

Substituting the expressions in (8) into (9), we obtain, after some simple manipulations, for $j < L$, the following relationship, for $j=1, \dots, L-1$:

$$\left(\frac{\partial \rho_E^j}{\partial P_E^j} \right)^{-1} \cdot (1 - \rho_E^j) = \left(\frac{\partial \rho_E^L}{\partial P_E^L} \right)^{-1} \cdot (1 - \rho_E^L) \cdot \frac{B_E^L}{B_E^j} \cdot \left[\sum_{l=j+1}^L \prod_{h=l}^L (1 - \rho_E^h)^{-1} \cdot \frac{\Delta_{l-1}}{\Delta_L} + 1 \right]. \quad (10)$$

In the final expression, the left side represents the information related to the j -th packet which depends only on the power of the following packets, $(j+1)$ to L . This observation motivates our recursive power allocation algorithm described in the following sections.

4 CHANNEL MODEL DEFINITION

We assume that each transmitted packet either arrives error free or is dropped due to a channel fade. If we define the channel model as in [12]:

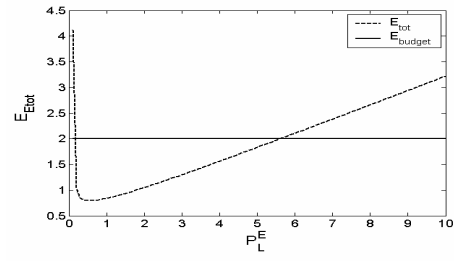
$$\rho_E^j = 1 - e^{-k/P_E^j}, \quad (11)$$

we can rearrange (10) and obtain the following expression:

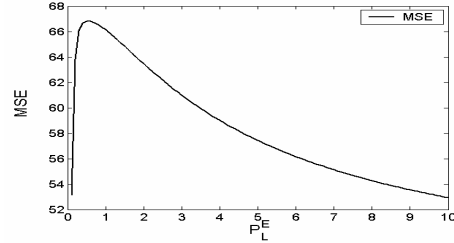
$$(P_E^j)^2 = (P_E^L)^2 \cdot \frac{B_E^L}{B_E^j} \cdot \left[\sum_{l=j+1}^L \prod_{i=l}^L e^{+k/P_E^i} \cdot \frac{\Delta_{l-1}}{\Delta_L} + 1 \right]. \quad (12)$$

The last expression allows us to recursively calculate all P_E^j ($j=1, \dots, L-1$) once P_E^L is known. In other words, for a given value of P_E^L we can recursively calculate P_E^j in terms of $\{P_E^{j+1}, \dots, P_E^L\}$, and therefore use (1) in order to calculate the resulting total energy. The minimization problem can then be solved by finding the value of P_E^L that satisfies the energy constraint. A closed form solution to the problem of finding the optimal P_E^L is difficult to compute analytically. Therefore, a numerical method such as the bisection method can be used.

Fig.2a shows, for a specific sequence, the plot of E_{Etot} versus the power of the last packet, P_E^L . As we can see from Fig.2a, for each value of E_{Etot} (i.e. for each energy budget) we can find two solutions (i.e., two points in which the derivative of the cost function equals zero): the algorithm must pick the one that gives the minimum distortion. Fig.2b shows the plot of the resulting MSE versus P_E^L .



(a)



(b)

Fig. 2 – Foreman QCIF sequence, frame 10, equal size packets: (a) total energy E_{Etot} and (b) frame distortion (MSE) versus the power of the last packet, P_E^L . In (a), the dotted line represents the energy budget.

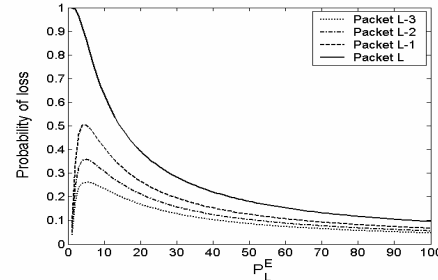


Fig.3 – Probability of packet loss ρ_E^j versus P_E^L for the last 4 packets ($j=L-3, \dots, L$).

If the line representing the energy constraint does not intersect with the energy curve, it means that a solution using a maximum number of packets $L_{max}=L$ does not exist, because all the solutions lead to a total energy that is larger than the energy constraint. In this case we have to consider the solution $P_E^L = 0$ (the last packet is not transmitted) and reconsider the problem by calculating the curve with $L_{max}=L-1$.

From the obtained results, we can observe that, in the case of equal size packets the power P_E^j assigned to each packet decreases with j : $P_E^j \geq P_E^{j+1}$. As a proof one can rearrange equation (12) in order to obtain:

$$(P_E^j)^2 = \frac{B_E^{j+1}}{B_E^j} \cdot (P_E^{j+1})^2 + \frac{B_E^L}{B_E^j} \cdot (P_E^L)^2 \cdot \frac{\Delta_j}{\Delta_L} \cdot \prod_{i=j+1}^L e^{+k/P_E^i}, \quad (13)$$

from which we can derive the following inequality:

$$P_E^j \geq \sqrt{\frac{B_E^{j+1}}{B_E^j}} \cdot P_E^{j+1}. \quad (14)$$

Since the power assigned to each packet is related by (11) to the probability of packet lost, this means that the probability of loss increases as the packet number j increases. The last packet (L -th) will be the packet with the lowest power, and maximum probability of loss (Fig.3).

In the dual problem, we want to minimize the total energy used for transmitting the enhancement layer, subject to a distortion constraint. In this case, the minimization problem can be expressed in the following way:

$$\min_{P_E^l} \left\{ \sum_{l=1}^L \frac{B_E^l \cdot P_E^l}{R} \right\}, \text{ s.t. } D_{E_{tot}} = D_B - E[\Delta_E] \quad (15)$$

where $D_{E_{tot}}$ is the maximum acceptable distortion for the frame. The techniques used to solve the minimum distortion problem (Eq. (2)) can also be used to solve the minimum energy problem (Eq. (15)). Note that for the minimum energy problem P_E^l is iteratively adjusted until the distortion constraint is met. The dual problem presented here is useful for applications in which a desired level of visual quality must be maintained using the least amount of energy.

6 EXPERIMENTAL SETUP AND RESULTS

In order to verify the proposed method, experiments with different types of packetization and coding parameters have been performed. The sequences have been coded using the MPEG-4 FGS reference software [13]. The Base Layer was coded with a fixed target bitrate R_B , utilizing the implementation of the TMN5 rate control present in the reference software. The motivation behind selecting a fixed R_B is that typically the Base Layer is compressed at a maximum bitrate R_B , such that it can always be received with negligible probability of error. The Enhancement Layer can then be partitioned in such a way that it can be transmitted whenever the bitrate is greater than R_B (and less than a certain R_E), fully utilizing the bandwidth available at the time of transmission. In this way, it is possible to adapt the coded video to the time-varying characteristics of the channel. Once coded, the EL is packetized using a fixed packet length scheme. Table 1 summarizes the parameters chosen for these experiments.

All the experiments reported here have been performed for the QCIF format (144x176 pixels) *Foreman* test sequence.

Similar results not presented in this paper due to space limitations have been obtained with other test sequences, such as *Carphone*, *Akiyo*, and *Salesman*. Three sets of experiments, named A, B and C, have been performed for analyzing the behavior of the algorithm, with different packet sizes and transmission rates (Table 2).

Table 1 – General parameter settings

E_{tot} for each frame	4 J
Frame rate	10 fps
Target bit rate for BL (R_B)	14 Kbps
Initial QP for I-Frame	20
Initial QP for P-Frame	30

Table 2 – Parameter settings for the three experiments

	Packet Size	Target Total Bit Rate
Experiment A	100 bytes	100 Kbps
Experiment B	100 bytes	400 Kbps
Experiment C	200 bytes	400 Kbps

In order to compare the results obtained from the proposed algorithm, an equal energy scheme and an empiric model have been implemented. The equal energy scheme assigns to each packet the same amount of energy. The empirical method varies the power per packet in a heuristic manor, and has been chosen in order to perform near the optimal algorithm. In experiment A, the empirical method assigns the power per packet in the following way:

$$P_E^k = \frac{E_{E_{tot}} \cdot R}{2^k \cdot B_E^k}, \quad k = 1, \dots, 7 \quad (16)$$

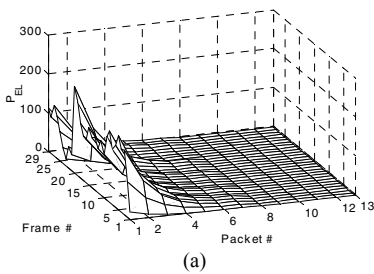
$$P_E^k = 0, \quad k > 7$$

In Fig. 4a and b, we plot, for each frame, respectively, the power assigned to each packet by the empirical method and the optimal power for experiment A.

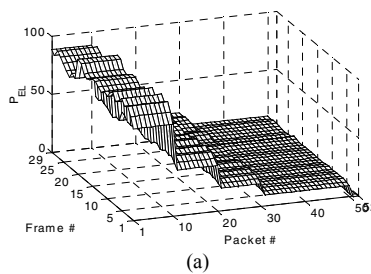
For experiment B the empirical method assigns the power per packet as follows:

$$P_E^k = \frac{E_{E_{tot}} \cdot R}{2^m \cdot \sum_{i=10 \times m - 9}^{10 \times m} B_E^i}, \quad 10 \times m - 9 < k \leq 10 \times m, \quad m \leq 5 \quad (17)$$

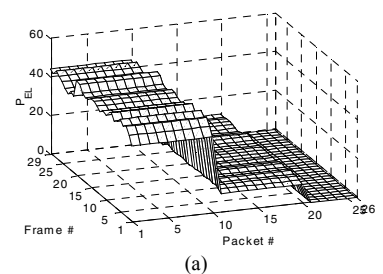
$$P_E^k = 0, \quad k > 50$$



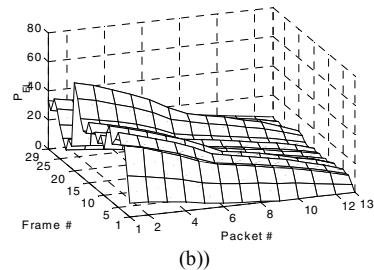
(a)



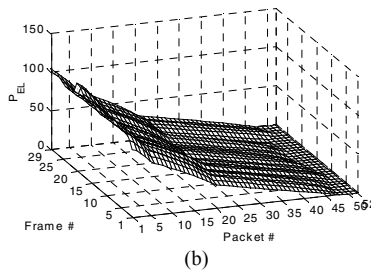
(a)



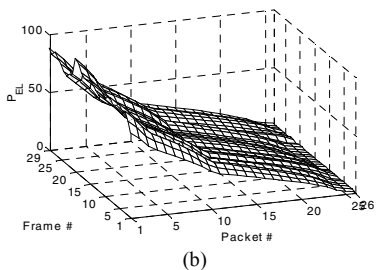
(a)



(b)



(b)

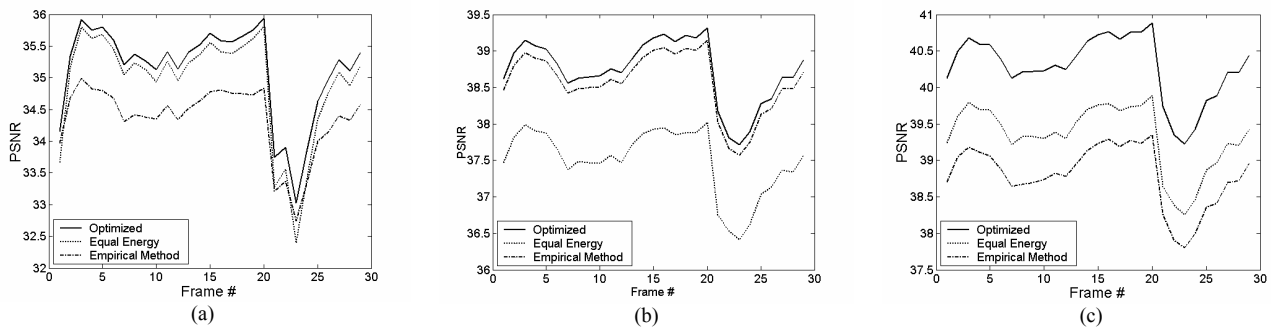


(b)

Fig.4 – Experiment A: (a) power per packet with the empirical method and (b) optimal power per packet

Fig.5 – Experiment B: (a) power per packet with the empirical method and (b) optimal power per packet

Fig.6 – Experiment C: (a) power per packet with the empirical method and (b) optimal power per packet



Fi.7 – PSNR: (a) experiment A, (b) B and (c) C

In Fig. 5a and b, we plot for each frame respectively the power assigned to each packet by the empirical method and the optimal power for experiment B.

Finally, for experiment C, the empirical method assigns power as follows:

$$P_E^k = \frac{E_{Etot} \cdot R}{2^m \cdot \sum_{i=5 \times m - 4}^{5 \times m} B_E^i}, 5 \times m - 4 < k \leq 5 \times m, m \leq 4 \quad (18)$$

$$P_E^k = 0, \quad k > 20$$

In Fig. 6a and b, we plot for each frame respectively the power assigned to each packet by the empirical method and the optimal power for experiment C.

The PSNR per frame for the proposed optimal power allocation approach, the equal energy scheme, and the empirical method are shown for experiments A, B, and C in Fig. 7. As expected, the PSNR obtained from the optimized method is always higher than both the equal energy and empirical methods. Fig. 7 demonstrates the potential of utilizing power adaptation to achieve unequal error protection in wireless video communications using FGS. As shown in Fig. 7, the optimized power allocation approach achieves an average expected PSNR that is on average 0.24 dB, 1.24 dB, and 0.95 dB higher than the equal energy approach for experiments A, B, and C respectively. The proposed approach is also able to achieve significant gains over the empirical method, which adapts the energy per packet in a heuristic manner. For example, in Fig. 7c, the empirical method achieves an average expected PSNR that is on average 1.48 dB lower than the optimized approach.

The performance of the equal power and empirical approaches strongly depends on the parameter settings as well as on the characteristics of the video sequence. Therefore, it is difficult to design a generic empirical algorithm that can always perform near the optimal power allocation. As shown in Fig. 7, an equal power per packet approach may perform near the optimized approach for some parameter setting (Fig. 7a), while the empirical approach may give a better approximation at other operating points (Fig. 7b). The proposed optimal power allocation scheme on the other hand is able to adaptively allocate the available energy based on the transmission parameters and the source characteristics.

7 CONCLUSIONS

An algorithm for optimally allocating the available transmission energy to the enhancement layer packets in an FGS wireless video communication system has been

presented. Unequal error protection is achieved through adaptive power allocation. The algorithm is simple, and can be used to efficiently calculate the optimal power distribution between the enhancement layer packets as well for validating faster empirical methods. The methodology used is applicable not only to FGS video coding but also other kinds of progressive coders, such as wavelet based image or video coding.

REFERENCES

- [1] ISO/IEC MPEG-4, "Information Technology – Coding of Audio Video Object – Part2: Visual – Amendment 4: Streaming Video Profile," MPEG 2000/N3518, July 2000.
- [2] W. Li, "Overview of fine granularity scalability in MPEG-4 video standard," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 11, pp. 301-317, March 2001.
- [3] M. van der Schaar, L. G. Boland, and Q. Li, "Novel applications of fine-granular-scalability: Internet & wireless video, scalable storage, personalized TV, universal media coding," *Proc. of SCI2001/ISAS2001 - Orlando, USA 2001*.
- [4] U. Horn, K. Stuhlmüller, M. Link, and B. Girod, "Robust Internet video transmission based on scalable coding and unequal error protection," *Image Communication*, vol. 15, no. 1-2, pp. 77-94, Sept. 1999.
- [5] M. van der Schaar and H. Radha, "Unequal packet loss resilience for Fine-Granular-Scalability video," *IEEE Trans. on Multimedia*, vol. 3, pp. 381-394, Dec. 2001.
- [6] X. K. Yang, C. Zhu, Z. G. Li, G. N. Feng, S. Wu, and N. Ling, "A degressive error protection algorithm for MPEG-4 FGS video streaming," *Proc. of IEEE ICIP*, Sept. 2002.
- [7] G. Wang, Q. Zhang, W. Zhu, and Y.-Q. Zhang, "Channel-adaptive unequal error protection for scalable video transmission over wireless channel," *Proc. of SPIE, Visual Communications and Image Processing*, pp. 648-655, San Jose, CA, Jan. 2001.
- [8] Q. Zhang, W. Zhu, and Y.-Q. Zhang, "Network-adaptive scalable video streaming over 3G wireless network," *Proc. IEEE ICIP, Thessaloniki, Greece, Oct. 2001*.
- [9] S. Zhao, Z. Xiong, and X. Wang, "Joint error control and power allocation for video transmission over CDMA networks with multiuser detection," *IEEE Trans. Circuits and Systems for Video Technology*, vol.12, pp. 425-437, June 2002.
- [10] Y. S. Chan and J. W. Modestino, "Transport of scalable video over CDMA wireless networks: A joint source coding and power control approach," *Proc. IEEE ICIP*, Oct. 2001.
- [11] Y. Eisenberg, C. E. Luna, T.N. Pappas, R. Berry, and A.K. Katsaggelos, "Joint source coding and transmission power management for energy efficient wireless video communications," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 12, pp. 411-424, June 2002.
- [12] L. Ozarow, S. Shamai, and A. Wyner, "Information theoretic considerations for cellular mobile radio," *IEEE Trans. Vehicular Tech.*, vol. 43, Issue 2, pp. 359-378, May 1994.
- [13] ISO/IEC MPEG-4 FGS Reference Software, available via <http://megaera.ee.nctu.edu.tw/mpeg/>, March 2003.