

Energy Efficient Video Summarization and Transmission Over a Slow Fading Wireless Channel

^{1,2}Zhu Li*, ^{2,3}Fan Zhai, ²Aggelos K. Katsaggelos, ²Thrasos N. Pappas

¹Multimedia Research Lab (MRL), Motorola Labs, Schaumburg, IL 60196

²Department of Electrical & Computer Engineering, Northwestern University, Evanston, IL 60208

³Digital Video Department, HPA, Texas Instruments, Dallas, TX 75243

ABSTRACT

With the deployment of 2.5G/3G cellular network infrastructure and large number of camera equipped cell phones, the demand for video enabled applications are high. However, for an uplink wireless channel, both the bandwidth and battery energy capability are limited in a mobile phone for the video communication. These technical problems need to be effectively addressed before the practical and affordable video applications can be made available to consumers. In this paper we investigate the energy efficient video communication solution through joint video summarization and transmission adaptation over a slow fading channel. Coding and modulation schemes, as well as packet transmission strategy are optimized and adapted to the unique packet arrival and delay characteristics of the video summaries. Operational energy efficiency – summary distortion performance is characterized under an optimal summarization setting.

Keywords: video summarization, wireless video, energy efficiency.

1. INTRODUCTION

With the proliferation of camera equipped cell phones and deployment of the higher data rate 2.5G and 3G infrastructure systems, providing video capability for consumers over the cellular communication infrastructure can drive a large number of valuable applications. However, wireless video communication presents some unique challenges. Due to shadowing and multi-path effect, the channel gain varies over time and this makes the reliable signaling difficult. On the other hand, communication of information over a wireless link consumes energy. For mobile devices, the energy is always limited, and even irreplaceable. How to achieve reliable video communication over a fading channel with energy efficiency is crucial for wide deployment of wireless video based applications.

There are quite a number of works in energy efficient wireless communication. A rather intuitive solution is to put the device into sleeping mode when not used. This has been explored in [12] and [14]. However, as the semiconductor technology improves, the energy consumption on circuits side can always be driving down, while the communication energy cost is lower bounded by information theory results, and its efficient usage need to be addressed. In [1], the fundamental trade-off between average power and delay constraint in communication over fading channels is explored and characterized. In [2], optimal power control schemes for communication over fading channels are developed. In [22] and [7], optimal off-line and near optimal on-line packet scheduling algorithms are developed to directly minimize energy used in communication a given amount of information over fading channels with certain delay constraints.

For streaming wireless video, the packet arrival pattern is not entirely random, and the delay constraint has certain unique aspects that differ it from other applications. Taking advantage of the specific source characteristics of video, and jointly adapting video source coding decisions with transmission power, modulation and coding schemes can achieve substantial energy efficiency compared with non-adaptive transmission schemes. Examples of this type of work are reported in [3], [6], [11], [13], and [18]. In those works, source-coding controls are mostly based on frame and / or macro block (MB) level coding mode and parameter decisions.

* Contact info: 1301 E Algonquin Road, IL02-Room2811, Schaumburg, IL 60196, USA. email: zhu.li@motorola.com; phone: 847 576-6942;

When both bandwidth and energy are severely limited for video streaming, sending over a video sequence with severe PSNR distortion is not desirable. Instead, we are looking at joint video summarization and transmission approaches to achieve energy efficiency. Video summarization has been an active research topic for a number of years. The main motivation for summarization is view time constraint. A shorter version of the original video sequence may be desirable in some applications, for example, a 2 minute summary of a 30 minute surveillance video can save people time in reviewing the video. Examples of the video summarization and related shot segmentation work can be found in [25], [5], [8], [21] and [17], [9], in which a video sequence is segmented into video shots, and then one or multiple key frames per shot are selected based on certain heuristic or metric for the summary.

Since the summarization process inevitably introduces distortion, and the summarization "rate" is related to the conciseness of the summary, we formulated the summarization problem as an rate-distortion optimization problem in [15], and developed an optimal solution based on dynamic programming. We extended the formulation to handle the case where bit rate is used as summarization rate in [16]. In this paper, we formulate the energy efficient video summarization and transmission problem as an energy-summarization distortion optimization problem, the solution is found through jointly optimizing the summarization and transmission parameters/decisions to achieve the operational optimality in energy efficiency performance.

The paper is organized into the following sections: In section 2 we give assumptions on the communication over fading channels and formulate the problem as an energy-summarization distortion optimization problem. In section 3 we develop an optimal solution based on Lagrangian relaxation and dynamic programming. In section 4 we present simulation results. Finally in section 5 we draw conclusions and discuss future works in this area.

2. ASSUMPTIONS AND PROBLEM FORMULATION

In this section we will discuss the channel model, delay analysis for video summary packets and the problem formulation.

2.1. Slow Fading Channel Models and Assumptions

Communication of video over a wireless channel has its unique challenges, namely fading and energy efficiency issues. In this work, we assume that the wireless channel can be modeled as a band-limited, AWGN channel with discrete time, slow block fading,

$$y_k = \sqrt{h_k} x_k + n_k \quad (1)$$

in which h_k is the channel gain for time slot k and it stays constant for time T_c , the channel coherent time. The channel is slow fading because we assume the symbol duration $T_s \ll T_c$. Thus, there are many channel uses during each time slot. n_k is the additive Gaussian noise with power N . The variation of the channel state is modeled as a finite state Markov channel (FSMC) [24], which has a finite set of possible states, $H = \{h_1, h_2, \dots, h_m\}$, and makes transition every T_c second with probability given by the transition probability matrix A , where $a_{ij} = \text{Prob}\{\text{transition from } h_j \text{ to } h_i\}$.

To reliably send R bits information over the fading channel in one channel use, the minimum power needed with optimal coding is given in information theory [4] as,

$$P = N(2^{2R} - 1)/h_k. \quad (2)$$

Similar to [Elif02], let $x=1/R$ be the number of transmission needed to send one bit over the channel, we can characterize the energy-delay tradeoff as E_b , energy per bit, as a function of x as,

$$E_b(x, h) = xP = xN(2^{2/x} - 1)/h \quad (3)$$

Examples of the energy efficiency functions with different fading states are shown in Fig. 1. The range of x in Fig. 1 corresponds to signal-to-noise ratio of 2.0dB to 20 dB, a typical operating range for wireless communication. To send a data packet with B bits and deadline τ , assuming $\tau \gg T_c$, the number of transmission available is $2W\tau$, where W is the signaling rate. Then the expected energy cost will be,

$$E(B, \tau) = E_H \{E_b(2W\tau/B, h)B | A, H, h_0\}. \quad (4)$$

In Eq. (4), the expectation is with respect to all possible channel states, which are governed by an FSMC specified by the state set H , the transition probability matrix A , and the initial state h_0 . The function in (4) can be implemented as a look up table for a given channel model in simulation. A closed form solution may also be possible, under some optimal coding and packet scheduling assumption.

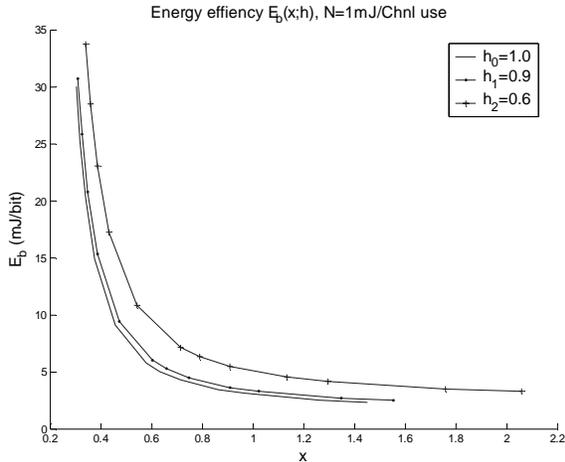


Figure 1. Energy efficiency over fading channels

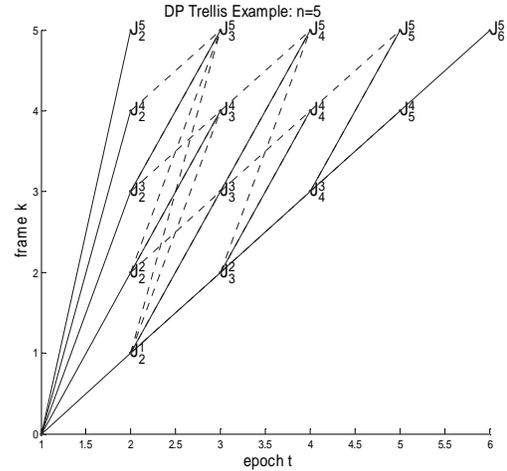


Figure 2. Example of DP Trellis

2.2. Summarization and Packet Delay Constraint Analysis

Let a video sequence of n frames be denoted by $V=\{f_0, f_1, \dots, f_{n-1}\}$, and its video summary of m frames be $S=\{f_{l_0}, f_{l_1}, \dots, f_{l_{m-1}}\}$. Obviously the video summarization process has an implicit constraint that $l_0 < l_1 < \dots < l_{m-1}$. Let the reconstructed sequence $V_S'=\{f_0', f_1', \dots, f_{n-1}'\}$ be obtained by substituting missing frames with the most recent frame that is in the summary S , that is,

$$f_k' = f_{i=\max(l): s.t. l \in \{l_0, l_1, \dots, l_{m-1}\}, i \leq k}.$$

The summarization distortion is therefore computed as the average frame distortion between the original sequence and the reconstructed sequence from the summary,

$$D(S) = \frac{1}{n} \sum_{k=0}^{n-1} d(f_k, f_k'). \quad (5)$$

where $d(f_k, f_k')$ is the distortion of the reconstructed frame f_k' , and n is the number of frames in the video sequence. For the encoding of the video summary frames, we assume a constant PSNR intra-coding strategy, with frame bit budget B_{l_j} given by some rate profiler, for example [10]. Packets from different summary frames have different delay tolerances. We assume that the first frame of the original sequence, f_0 , is always selected for the summary and intra coded with some B_0 bits. The delay tolerance τ_0 is determined by how much initial streaming delay is allowed. For packets generated by the summary frame f_{l_j} , with $l_j > 0$, if the previous summary frame $f_{l_{j-1}}$ is decoded at time t_{j-1} , then the packet need to arrive by the time $t_j=t_{j-1} + (l_j - l_{j-1})/F$, where F is the frame rate of the sequence. Therefore, the delay tolerance for frame f_{l_j} is $\tau_{l_j}=(l_j - l_{j-1})/F$. This is a simplified delay model, not accounting for minor variations in frame encoding and other delays. The energy cost to transmit a summary S of m frames is therefore given by,

$$E(S) = \sum_{k=0}^{m-1} E(B_{l_k}, \tau_{l_k}) = E(B_0, \tau_0) + \sum_{k=1}^{m-1} E(B_{l_k}, \tau_{l_k}) \quad (6)$$

where B_{l_k} is the number of bits needed to encode summary frame f_{l_k} , and τ_{l_k} is the delay tolerance for frame f_{l_k} packets. There are trade offs between the number of frames, m , in the summary versus the summarization distortion. The more frames can be selected into the summary, the smaller the summarization distortion. On the other hand, the more frames in the summary, more bits need to be spend encoding the frames, and the packet arrival pattern gets more

crowded, which translate into higher bit rate and smaller delay tolerance. More bits to transmit with more stringent deadline translate into higher transmission energy cost. Let the summarization rate be,

$$R(S) = \frac{m}{n}$$

In the next sub-section, we will characterize the relationship between the summarization distortion and energy cost, and formulate the energy efficient video summarization and transmission problem as an Energy-Distortion (ED) optimization problem.

2.3. Energy Efficient Summarization Formulations

The energy efficient summarization problem can be formulated as a constrained optimization problem. For a given constraint on the summarization distortion, we need to find the optimal summary that minimizes the transmission energy cost, while satisfying the distortion constraint, D_{max} , that is, the Minimizing Energy Optimal Summarization (MEOS) formulation,

$$S^* = \arg \min_S E(S), \text{ s.t. } D(S) \leq D_{max}. \quad (7)$$

We can also formulate the energy efficiency problem as the Minimizing Distortion Optimal Summarization (MDOS) problem, that is, for a given energy constraint, E_{max} , find the optimal summary that minimizes the summarization distortion,

$$S^* = \arg \min_S D(S), \text{ s.t. } E(S) \leq E_{max}. \quad (8)$$

The solution to the formulations in Eqs. (7) and (8) is based on Lagrangian relaxation and Dynamic Programming (DP) which is given in the next section.

3. OPTIMAL SOLUTION

Directly solving the constrained problems in Eqs. (7) and (8) is usually difficult due to the complicated dependencies and large searching space for the operating parameters. Instead, we introduce the Lagrangian multiplier relaxation to convert the original problem into an unconstrained problem. The solution to the original problem can then be found by solving the resulting unconstrained problem with the appropriate Lagrangian multiplier that satisfies the constraint. This approach has been widely used in solving a number of coding and resource allocation problems in video/image compression [20].

Considering the MEOS formulation, by introducing the Lagrangian multiplier, the relaxed problem is therefore given by,

$$S^*(\lambda) = \arg \min_S \{E(S) + \lambda D(S)\}, \quad (9)$$

in which the optimal solution S^* becomes a function of λ . From [19] we know that by varying λ from zero to infinity, we sweep the convex hull of the operational Energy-Distortion (ED) function $E(D(S^*(\lambda)))$, which is also monotonic with respect to λ . Therefore a bi-section search algorithm on λ can give us the optimal solution within a convex hull approximation. In real world applications, the operational ED functions are typically convex, and indeed the optimal solution can be found by the algorithm above.

To solve the relaxed problem in (9) by exhaustive search, which has an exponential computational complexity, is not feasible. Instead, we observe that there are built-in recursive structures that can be exploited for an efficient dynamic programming solution for the relaxed problem with polynomial computational complexity.

First, let us introduce a notation on segment distortion for distortion introduced by missing frames between summary frame l_t and l_{t+1} , which is given by,

$$G_{l_t}^{l_{t+1}} = \sum_{k=l_t}^{l_{t+1}-1} d(f_l, f_k). \quad (10)$$

Let the “state” of a video summary with t -frames, and the last frame f_k be the minimum of the relaxed objective function be given by,

$$\begin{aligned}
J_t^k(\lambda) &= \min_{S:s.t.,|S|=t,l_{t-1}=k} \{D(S) + \lambda E(S)\} \\
&= \min_{l_1, l_2, \dots, l_{t-2}} \{G_0^{l_1} + G_{l_1}^{l_2} + \dots + G_{l_{t-2}}^k + G_k^n + \lambda \sum_{k=0}^{t-1} E(B_{l_k}, \tau_{l_k})\},
\end{aligned} \tag{11}$$

where $|S|$ denotes the number of frames in S . Notice that $l_0=0$, as we assume the first frame is always selected. The minimization process in Eq. (10) has the following recursion,

$$\begin{aligned}
J_{t+1}^k(\lambda) &= \min_{S:s.t.,|S|=t+1,l_t=k} \{D(S) + \lambda E(S)\} \\
&= \min_{l_1, l_2, \dots, l_{t-1}} \{G_0^{l_1} + G_{l_1}^{l_2} \dots + G_{l_{t-1}}^k + G_k^n + \\
&\quad \lambda [E(B_0, \tau_0) + E(B_{l_1}, (l_1 - 0)/F) + \dots + E(B_{l_{t-1}}, (l_{t-1} - l_{t-2})/F) + E(B_k, (k - l_{t-1})/F)]\} \\
&= \min_{l_1, l_2, \dots, l_{t-1}} \underbrace{\{G_0^{l_1} + G_{l_1}^{l_2} \dots + G_{l_{t-1}}^{l_{t-1}} + G_{l_{t-1}}^n - G_{l_{t-1}}^n + G_{l_{t-1}}^k + G_k^n + \\
&\quad \lambda [E(B_0, \tau_0) + E(B_{l_1}, (l_1 - 0)/F) + \dots + E(B_{l_{t-1}}, (l_{t-1} - l_{t-2})/F) + E(B_k, (k - l_{t-1})/F)]\}}_{D_t^{l_{t-1}}} \\
&= \min_{l_1, l_2, \dots, l_{t-1}} \underbrace{\{D_t^{l_{t-1}} + \lambda E_t^{l_{t-1}} + \lambda E(B_k, (k - l_{t-1})/F) - G_{l_{t-1}}^n + G_{l_{t-1}}^k + G_k^n\}}_{e^{l_{t-1}, k}} \\
&= \min_{l_{t-1}} \{J_t^{l_{t-1}}(\lambda) + e^{l_{t-1}, k}\}
\end{aligned} \tag{12}$$

The recursion has the initial condition given by,

$$J_1^0(\lambda) = G_0^n + \lambda E(B_0, \tau_0). \tag{13}$$

The Dynamic Program (DP) solution starts with the initial node J_1^0 , and propagates through a trellis with arcs representing possible transitions. The cost of transition is given by edge cost $e^{l_{t-1}, k}$ in Eq. (12). At each node we compute and store the optimal incoming arc and minimum cost. Once all the nodes with final virtual frame f_n , $\{J_t^n(\lambda) | t = 1, 2, \dots, n\}$, are computed, the optimal solution to the relaxed problem in Eq. (9) is found by selecting the minimum cost final virtual frame nodes,

$$S^*(\lambda) = \arg \min_t \{J_t^n(\lambda)\},$$

and backtracking for the optimal solution. This is similar to the Viterbi algorithm [23]. An example trellis for $n=5$ is shown in Fig. 2. All possible state transitions are plotted. For each state node, the minimum cost incoming is plotted in solid line, while other incoming arcs are plotted as dotted lines. For example, the node J_3^4 is computed as $J_3^4 = \min_{j \in \{1, 2, 3\}} \{J_2^j + e^{j, 3}\}$, and its minimum cost incoming arc is from node J_2^2 . The virtual final frame nodes are all at the top of the trellis.

The Lagrangian multiplier controls the trade off between the summarization distortion and the energy cost in transmitting the summary. By varying the value of λ and solving the relaxed problem in the inner loop, we can obtain the optimal solution that minimizing the transmission energy cost while meeting certain distortion constraints. Since the operational energy-distortion function $E(D(S^*(\lambda)))$ is monotonic with λ , a fast bi-section search algorithm can be applied to find the optimal λ^* , that results in the tightest bound on the distortion constraint D_{max} , i.e., $D(S^*(\lambda^*))$ is closest to D_{max} . The algorithm can be made even faster by re-using the distortion and energy cost results that only need to be computed once in the iteration. The solution to the MDOS formulation can also be solved in the same fashion.

4. SIMULATION RESULTS

To simulate a slow fading wireless channel, we model the channel fading as a two state FSMC with channel states h_0 and h_1 . The channel state transitional probability is given by $A = \begin{bmatrix} 1-p & p \\ q & 1-q \end{bmatrix}$, and is shown in Fig. 3. In Fig. 3, p and q are the probabilities of channel state transition from h_0 to h_1 , and h_1 to h_0 , respectively. The steady state channel state probability is therefore computed as $\pi_0 = \frac{q}{p+q}$ and $\pi_1 = \frac{p}{p+q}$. Assuming that $\tau \gg T_c$, and the signaling rate is W (W is selected to simulate typical SNR operating range in wireless communications), then out of total $2W\tau$ channel uses, $\frac{p}{p+q} 2W\tau$ of them are in channel state h_1 , and $\frac{q}{p+q} 2W\tau$ of them are in channel state h_0 .

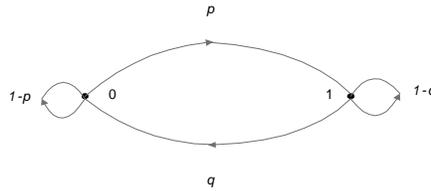


Figure 3. Two state FSMC example

Assuming the channel state is known to both the transmitter and the receiver, with the optimal coding and packet scheduling, then the expected energy cost of transmitting B bits with delay constraint τ can then be computed as,

$$\begin{aligned}
 E(B, \tau) &= E_H \{ E_b(2W\tau/B, h)B \} \\
 &= \min_{0 \leq x \leq 1} \{ f(x; B, W, \tau, p, q, h_0, h_1) \} \\
 &= \min_{0 \leq x \leq 1} \{ xBE_b(\frac{q}{p+q} 2W\tau/(xB), h_0) + (1-x)BE_b(\frac{p}{p+q} 2W\tau/(B(1-x)), h_1) \}
 \end{aligned} \tag{14}$$

In Eq. (14), we need to find an optimal bits splitting, x , of the total bits B , with xB bits transmitted optimally while channel state is in h_0 , and $(1-x)B$ bits transmitted optimally while channel state is h_1 .

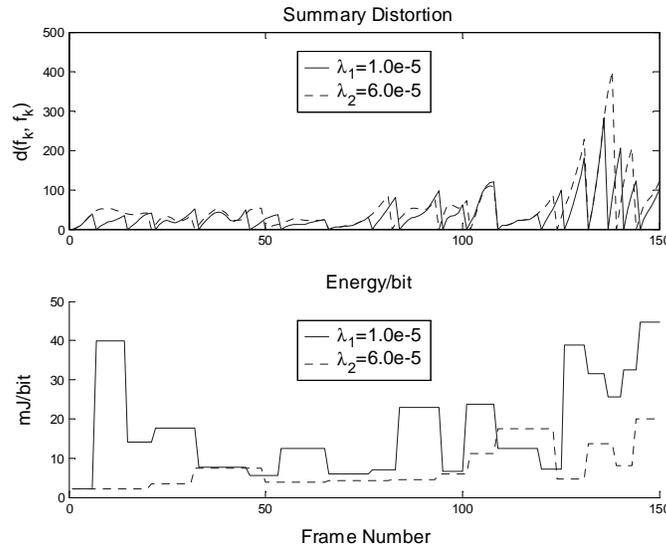


Figure 4. Examples of Energy Efficient Video Summarization

Notice that Eq. (14) can be implemented as a look up table in a practical system implementation. For simple channel models such as the two-state FSMC, a closed form solution can be derived. Checking the 1st and 2nd order conditions (See Appendix for more detail) for the minimization problem in Eq. (14), the optimal splitting of the bits is given by,

$$x^* = \frac{w\tau pq}{B(p+q)^2} [\log_2(\frac{h_0}{h_1}) + \frac{(p+q)}{w\tau} B] = \frac{w\tau pq}{B(p+q)^2} \log_2(\frac{h_0}{h_1}) + \frac{q}{(p+q)}, \quad (15)$$

and the minimum energy cost is given by,

$$E(B, \tau) = f(x^*; B, W, \tau, p, q, h_0, h_1) \\ = x^* BE_b(\frac{q}{p+q} 2W\tau/(x^* B), h_0) + (1-x^*) BE_b(\frac{p}{p+q} 2W\tau/(B(1-x^*)), h_1) \quad (16)$$

Eq. (16) can be implemented as a look up table for the ED optimization algorithm.

In our simulation, the optimal algorithm is implemented in Matlab and simulation results for the “foreman” sequence, frames 150~299 are shown in the Fig. 4. The channel state is modeled as $h_0=0.9, h_1=0.1, p=0.7, q=0.8$. Signaling rate is $W=20\text{kHz}$. The background noise power is assumed to be $N=1\text{mJ}$ per channel use. The summary frames are intra-coded with constant PSNR quality with H.263 implementation TMN5. Summarization distortion and average power during transmission are plotted for different values of Lagrangian multiplier, with $\lambda_1=1.0\text{e-}5$ and $\lambda_2=6.0\text{e-}5$. For the larger Lagrangian value, λ_2 , more emphasis is placed on minimizing the energy cost, therefore the energy cost (area under the average power plot) is smaller than the smaller value λ_1 . On the other hand, the summarization distortion is larger for λ_1 than λ_2 , as expected.

The overall operational E-D performances are characterized as the E-D curves in the Fig. 6 below for both $W=20\text{kHz}$ and $W=40\text{kHz}$ cases.

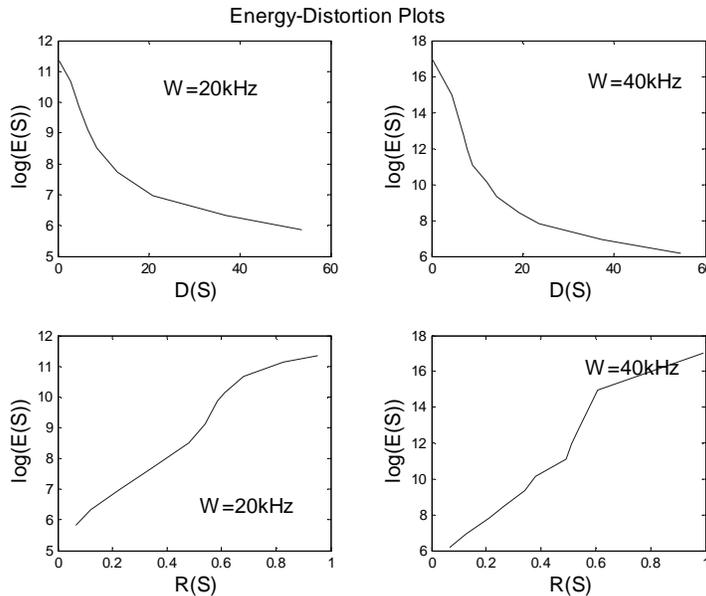


Figure 5 . Energy-Distortion Performance

For the same 150-frame segment of the “foreman” sequence, the upper two plots characterize the relationship between the summarization distortion and the total energy cost in $\log_{10}(\text{mJ})$ scale. As the summarization distortion goes up linearly, the energy cost drops exponentially. The lower two plots characterize the energy cost in terms of the summarization rate $R(S) = m/n$, where m is the number of frames in S . It is clear that in the typical operating range of the video summarization, for example, $R(S)=[0.2, 0.8]$, the communication energy cost can change in the order of 5 to 10 magnitudes. This indicates that summarization is a very efficient energy conserving scheme for wireless video communications.

5. CONCLUSION AND FUTURE WORK

In this work we proposed an optimal (within a convex hull approximation) algorithm for energy efficient video summarization and transmission. The algorithm is based on the Lagrangian relaxation and dynamic programming. The simulation results indicate that this is a very efficient and effective method in optimal communication energy utilization for video transmission over a fading channel.

In the future, we will extend the work with more complex channel simulations, and develop algorithms that allow inter-frame coding scheme for summaries. We will also investigate computation complexity reduction issues by looking for heuristic algorithms that can offer near-optimal performances.

REFERENCE

1. R. Berry and R. Gallager, "Communication over fading channels with delay constraints", *IEEE Trans. on Information Theory*, May, 2002.
2. G. Caire, G. Taricco, and E. Biglieri, "Optimum power control over fading channels", *IEEE Trans. on Information Theory*, July 1999.
3. Y. S. Chan and J. W. Modestino, "Transport of scalable video over CDMA wireless networks: A joint source coding and power control approach", *Proceedings of Int'l Conference on Image Processing (ICIP)*, Thessaloniki, Greece, 2001.
4. T. Cover and J. Thomas, *Elements of Information Theory*, Wiley Series in Telecommunication, New York, 1991.
5. N. Doulamis, A. Doulamis, Y. Avrithis and S. Kollias, "Video Content Representation Using Optimal Extraction of Frames and Scenes", *Proc. of Int'l Conference on Image Processing (ICIP)*, Chicago, Illinois, 1998.
6. Y. Eisenberg, C. E. Luna, T.N. Pappas, R. Berry, and A. K. Katsaggelos, "", *IEEE Trans. on Circuits & System for Video Technology*, vol. 12, no. 8, June 2002.
7. A. El. Gamal, C. Nair, B. Prabhakar, E. Uysal-biyikoglu, and S. Zahedi, "Energy efficient scheduling of packet transmission over wireless networks", *Proceedings of INFOCOM*, 2002.
8. A. Hanjalic and H. Zhang, "An Integrated Scheme for Automated Video Abstraction Based on Unsupervised Cluster-Validity Analysis", *IEEE Trans. on Circuits and Systems for Video Technology*, vol.9, December 1999.
9. A. Hanjalic, "Shot-Boundary Detection: Unraveled and Resolved?", *IEEE Trans. on Circuits and Systems for Video Technology*, vol.12, No. 2, February 2002.
10. Zhihai He and Sanjit K. Mitra, "A Unified Rate-Distortion Analysis Framework for Transform Coding," *IEEE Trans. on Circuits and System for Video Technology*, vol. 11, pp. 1221 –1236, December 2001
11. Z. He, Jianfei Cai, and Chang Wen Chen, "Joint source channel rate-distortion analysis for adaptive mode selection and rate control in wireless video coding", *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 12, no. 6, June 2002.
12. Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY), Spec. of IEEE 802.11 Standard, 1998.
13. Il-Min Kim, and Hyung-Myung Kim, "An optimum power management scheme for wireless video service in CDMA systems", *IEEE Trans. on Wireless Communication*, vol. 2, no. 1, January, 2003.
14. R. Kravets and P. Krishnan, "Application-driven power management for mobile communication", *Wireless Networks*, vol. 6, no. 4, September, 2000.
15. Z. Li, A. K. Katsaggelos, G. Schuster, and B. Gandhi, "Rate-Distortion Optimal Video Summary Generation", *IEEE Trans. on Image Processing*, to appear.
16. Z. Li, G. Schuster, A. K. Katsaggelos, and B. Gandhi, "Bit Constrained Optimal Video Summarization", *Proceedings of International Conference on Image Processing (ICIP)*, Singapore, 2004.
17. R. Lienhart, "Reliable Transition Detection in Videos: A Survey and Practitioner's Guide", *International Journal of Image and Graphics*, Vol.1, No.3, pp. 469-486, 2001.
18. C. E. Luna, Y. Eisenberg, R. Berry, T. N. Pappas, and A. K. Katsaggelos, "Joint Source Coding and Data Rate Adaptation for Energy Efficient Wireless Video Streaming", *IEEE Journal on Selected Areas in Communications*, vol.21, no. 10, December 2003.
19. K. Ramchandran and M. Vetterli, "Best wavelet packet bases in a rate-distortion sense", *IEEE Trans. on Image Processing*, vol. 2, no. 2, April 1993.

20. G. M. Schuster and A. K. Katsaggelos, Rate-Distortion Based Video Compression, Optimal Video Frame Compression and Object Boundary Encoding. Norwell, MA: Kluwer, 1997.
21. H. Sundaram and S-F. Chang, "Constrained Utility Maximization for Generating Visual Skims", *IEEE Workshop on Content-Based Access of Image & Video Library*, 2001.
22. E. Uysal-Biyikoglu, B. Prabhakar, and A. E. Gamal, "Energy Efficient Packet Transmission Over a Wireless Link", *IEEE/ACM Trans. on Networking*, vol. 10, no. 4, August 2002.
23. A. J. Viterbi, "Error Bounds for Convolutional Codes and an Asymptotically Optimum Decoding Algorithm", *IEEE Trans. on Information Theory*, vol. IT-13, pp. 260-269, April 1967.
24. H. S. Wang, and N. Moayeri, "Finite-state Markov channel- a useful model for radio communication channels", *IEEE Trans. on Vehicular Tech*, vol. 44, February 1995.
25. Y. Zhuang, Y. Rui, T. S. Huan, and S. Mehrotra, "Adaptive Key Frame Extracting Using Unsupervised Clustering", *Proc. of Int'l Conference on Image Processing (ICIP)*, Chicago, Illinois, 1998

6. APPENDIX

Derivation of the optimal bit split solution is given below:

Assuming the channel state is known to both the transmitter and the receiver, the expected energy cost of transmitting B bits with delay τ is computed as,

$$\begin{aligned}
 E(B, \tau) &= E_H \{E_b(2W\tau/B, h)B\} \\
 &= \min_{0 \leq x \leq 1} \{f(x; B, W, \tau, p, q, h_0, h_1)\} \\
 &= \min_{0 \leq x \leq 1} \{xBE_b(\frac{q}{p+q} 2W\tau/(xB), h_0) + (1-x)BE_b(\frac{p}{p+q} 2W\tau/(B(1-x)), h_1)\}
 \end{aligned}$$

So we have,

$$\begin{aligned}
 f(x) &= xBE_b(2W\tau\pi_0/(xB), h_0) + (1-x)BE_b(2W\tau\pi_1/((1-x)B), h_1) \\
 &= (2\pi_0W\tau/h_0)(2^{\frac{xB}{\pi_0W\tau}} - 1) + (2\pi_1W\tau/h_1)(2^{\frac{(1-x)B}{\pi_1W\tau}} - 1)
 \end{aligned}$$

Let,

$$\begin{aligned}
 a_0 &= 2\pi_0W\tau/h_0, & a_1 &= 2\pi_1W\tau/h_1, \\
 b_0 &= \frac{B}{\pi_0W\tau}, & b_1 &= \frac{B}{\pi_1W\tau}
 \end{aligned}$$

We have, $f(x) = a_0(2^{b_0x} - 1) + a_1(2^{b_1(1-x)} - 1)$. To minimize $f(x)$, let the 1st order condition be zero, this leads to,

$$f'(x) = a_0b_0 \ln(2)2^{b_0x} - a_1b_1 \ln(2)2^{b_1(1-x)} = 0, \Rightarrow x^* = \frac{1}{b_0 + b_1} (\log_2(\frac{a_1b_1}{a_0b_0}) + b_1),$$

and since the 2nd order condition is always non-negative:

$$f''(x) = a_0b_0^2 \ln^2(2)2^{b_0x} + a_1b_1^2 \ln^2(2)2^{b_1(1-x)} \geq 0, \forall 0 \leq x \leq 1,$$

the optimal bit splitting ratio is,

$$x^* = \pi_0\pi_1 \log_2\left(\frac{h_0}{h_1}\right) \frac{W\tau}{B} + \pi_0,$$

and the optimal energy cost is given by,

$$E(B, \tau) = x^* BE_b(2\pi_0W\tau/(x^*B), h_0) + (1-x^*)BE_b(2\pi_1W\tau/(B(1-x^*)), h_1)$$