Energy Minimization of Portable Video Communication Devices Based on Power-Rate-Distortion Optimization

Zhihai He, *Senior Member, IEEE*, Wenye Cheng, and Xi Chen

***Abstract*—Portable video communication devices operate on batteries with limited energy supply. However, video compression is computationally intensive and energy-demanding. Therefore, one of the central challenging issues in portable video communication system design is to minimize the energy consumption of video encoding so as to prolong the operational lifetime of portable video devices. In this work, based on power-rate-distortion (P-R-D) optimization, we develop a new approach for energy minimization by exploring the energy tradeoff between video encoding and wireless communication and exploiting the nonstationary characteristics of input video data. Both analytically and experimentally, we demonstrate that incorporating the third dimension of power consumption into conventional R-D analysis gives us one extra dimension of flexibility in resource allocation and allows us to achieve significant energy saving. Within the P-R-D analysis framework, power is tightly coupled with rate, enabling us to trade *bits* for *joules* and perform energy minimization through optimum bit allocation. We analyze the energy saving gain of P-R-D optimization. We develop an adaptive scheme to estimate P-R-D model parameters and perform online resource allocation and energy optimization for real-time video encoding. Our experimental studies show that, for typical videos with nonstationary scene statistics, using the proposed P-R-D optimization technology, the energy consumption of video encoding can be significantly reduced (by up to 50%), especially in delay-tolerant portable video communication applications.**

***Index Terms*—Lifetime maximization, power consumption, rate-distortion (R-D), resource allocation, video compression.**

I. Introduction

Wireless video communication over portable devices has become the driving technology of many important applications, experiencing dramatic market growth and promising revolutionary experiences in personal communication, gaming, entertainment, military, security, environment monitoring, and more [1], [2]. Portable devices are powered by batteries. Video encoding schemes are often computationally intensive and energy-demanding, even after being fully optimized with existing software and hardware energy-minimization techniques [3], [4]. As a result, the operational lifetime of most portable video systems, such as handheld video devices, still remains short, mostly in the range of a few hours. This has become a bottleneck for technological progress in portable video electronics.

Video data is voluminous. It has to be very efficiently compressed. Otherwise, the amount of transmission energy, transmission bandwidth, or required storage space[1] will be tremendous. During the past two decades, many video compression algorithms and international standards, such as MPEG-2, H.263, MPEG-4, and H.264 [5], [6], have been developed for efficient video compression. An efficient video compression system is often computationally intensive and energy-consuming, since it involves many sophisticated operations in spatiotemporal prediction, transform, quantization, mode selection, and entropy coding [7]. Recent studies [4], [8], as well as our experimental analysis to be described in Section III, show that in typical scenarios of video communication over portable devices video encoding consumes a significant portion (up to 40%–60%) of the total energy. Therefore, one of the important issues in portable video communication system design is to minimize the energy consumption of video encoding so as to prolong the operational lifetime of portable video devices.

A. Related Work

To reduce the energy consumption of video encoders, a number of algorithms, software and hardware energy-minimization techniques, including low-complexity encoder design [9]–[11], low-power embedded video encoding algorithms [12], adaptive power control methods [13], [4], [8], and joint encoder and hardware adaptation schemes [3], [14], [15] have been proposed. During the past decades, fast algorithms for major video encoding modules, including motion estimation, mode decision, and transform, have been developed [11], [16]–[19]. [20] provides a comprehensive review of fast algorithms for motion estimation in video compression. Hardware implementation technologies have also been developed to improve the video encoding speed [12], [21], [24], [25]. Recently, researchers have realized the importance of cross-layer design for energy saving in multimedia systems [3], [15]. Joint adaptation of video encoder/decoder and hardware scheduling for energy saving has been studied [3], [14], [15], [22]. Joint encoder and

[1]In many applications, such as personal video recording, storage space may still be a major system resource constraint.

Manuscript received April 6, 2007; revised August 28, 2007. This work was supported in part by the National Science Foundation under Grant DBI-0529082. This paper was recommended by Associate Editor D. S. Turaga.

Z. He and W. Cheng are with the Department of Electrical and Computer Engineering, University of Missouri, Columbia, MO 65211 USA (e-mail: HeZhi@missouri.edu).

X. Chen was with the Department of Electrical and Computer Engineering, University of Missouri, Columbia, MO 65211 USA. He is now with Harmonic Inc., Sunnyvale, CA 94089 USA.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCSVT.2008.918802

wireless transmission power control schemes have also been developed, for example [4].

So far, existing methods and algorithms have been focused on encoder complexity (and power consumption) reduction through heuristic adaptation or control instead of systematic energy optimization. This is because they lack an analytic model to characterize the optimum trade-off between energy consumption and encoding performance [7]. In addition, even with existing energy saving technologies, the operational lifetime of portable video electronics still remains very short, which has become one of the biggest impediments to our technology future. Therefore, it is important to develop new energy minimization approaches for portable video communication devices.

B. This Work

In this work, we study the following problem: given a portable video encoding system, whose power consumption has already been fully optimized with existing software and hardware techniques, can we develop a new approach to achieve additional energy saving and further extend its operational lifetime? Our answer is YES. Our approach is based on power-rate-distortion (P-R-D) optimization and control. In this work, we will propose an *operational* framework for analyzing and modeling the P-R-D behaviors of generic video encoders. We will demonstrate that extending the traditional R-D analysis to P-R-D analysis will give us another dimension of flexibility in resource allocation and performance optimization for wireless video communication over portable devices. We will see that, within the new P-R-D analysis framework, bit and energy resources are tightly coupled, which enables us to trade " *bits* " for " *joules* " (energy) so as to achieve significant energy saving for nonstationary video sources. We will analyze the energy saving gain of P-R-D optimization. We will develop an adaptive scheme to estimate the P-R-D model parameters and perform online energy optimization for real-time video compression. Our extensive experimental studies show that, for typical videos with nonstationary statistics, using the proposed P-R-D optimization technology, the encoder energy consumption can be significantly reduced. This has many important applications in low-power portable video communication system design.

The rest of the paper is organized as follows. The operational P-R-D analysis and analytic P-R-D models for generic video encoders are presented in Section II. In Section III, we discuss practical portable video communication system design, study its energy consumption behaviors, and explain how video encoding energy minimization can be performed within the context of wireless video communication over portable devices. In Section IV, we will study how video encoding energy minimization can be performed through optimum bit allocation and analyze the energy saving gain. Section V presents the online energy minimization algorithm for real-time video encoding. Experimental results are presented in Section VI. Section VII concludes the paper and discusses future research directions.

II. P-R-D Modeling of Video Encoders

In this section, we introduce a generic operational method to study the R-D behavior of video encoders under power constraints and establish their P-R-D models. The central task of P-R-D analysis is to answer the following question: *what is the minimum distortion that a video encoder can achieve under rate and power constraints*?

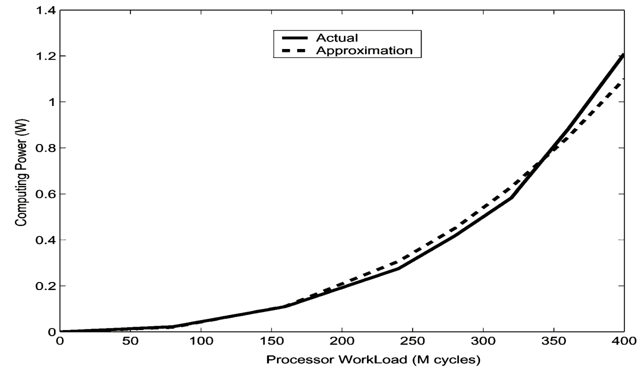


Fig. 1. Power consumption model with DVS.

In our previous work of P-R-D analysis [7], we have introduced a set of complexity control parameters into an MPEG-4 encoder, analyzed the R-D behavior of each parameter, and obtained the P-R-D function for a simple MPEG-4 encoder. We observe that this analytical approach cannot be easily extended to other video encoders, such as H.264 video coding [6]. Direct R-D analysis of complexity control parameters becomes very difficult when the video encoding mechanism becomes more sophisticated. In this work, we propose an *operational* approach for offline P-R-D analysis and modeling which can be applied to generic video encoders. By performing this offline P-R-D analysis procedure over a wide range of training video data, we will be able to establish models and a control database for online resource allocation and energy minimization.

A. Operational P-R-D Analysis

The operational P-R-D analysis has the following three major steps. In the **first** step, we group major encoding operations into several modules, such as motion prediction, pre-coding (transform and quantization), mode decision, and entropy coding, and then introduce a set of control parameters $\Gamma = [\gamma_1, \gamma_2, \ldots, \gamma_L]$ to control the computational complexity of these modules. Therefore, the encoder complexity (or processor workload) C is then a function of these control parameters, denoted by $C(\gamma_1, \gamma_2, \ldots, \gamma_L)$. Within the DVS (dynamic voltage scaling) design framework [23], the microprocessor power consumption, denoted by P, is a function of processor workload C, therefore, is also a function of Γ, denoted by

$$P = \Phi(C) = P(\gamma_1, \gamma_2, \ldots, \gamma_L) \tag{1}$$

where $\Phi(\cdot)$ is the power consumption model of the microprocessor [23]. For example, according to our measurement to be described in III, the power consumption model of the Intel PXA255 XScale processor is shown in Fig. 1 (solid line). It can be well approximated by the following expression

$$P = \Phi(C) = \beta \times C^{\gamma}, \quad \gamma = 2.5 \tag{2}$$

where β is a constant. In the **second** step, we execute the video encoder with different configurations of complexity control pa-

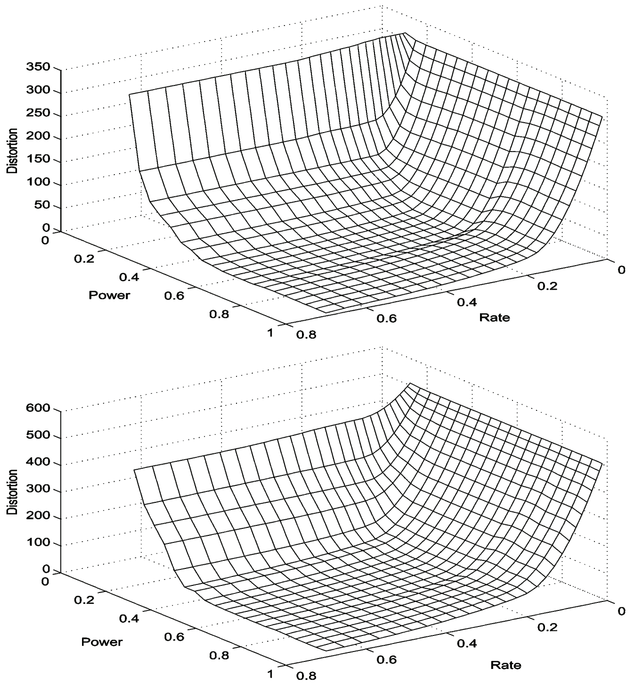


Fig. 2. Operational P-R-D functions of “Foreman” (top) and “Football” (bottom) CIF videos encoded with MPEG-4 at 10 fps.

rameters and obtain the corresponding R-D data, denoted by $D(R; \gamma_1, \gamma_2, \ldots, \gamma_L)$. Note that this step is computationally intensive and is intended for offline analysis to obtain the P-R-D model only. Once the model is established, in Section II -C, we will discuss how the model parameter can be estimated during online video encoding.

In the **third** step, we perform optimum configuration of the complexity control parameters to maximize the video quality (or minimize the video distortion) under the power constraint. This optimization problem can be mathematically formulated as follows:

$$\min_{\{\gamma_1, \gamma_2, \ldots, \gamma_L\}} \quad D = D(R; \gamma_1, \gamma_2, \ldots, \gamma_L) \\ \text{s.t.} \quad P(\gamma_1, \gamma_2, \ldots, \gamma_L) \leq P \tag{3}$$

where P is the available power consumption for video encoding. Given the R-D data set $\{D(R; \gamma_1, \gamma_2, \ldots, \gamma_L)\}$, this minimization problem can be easily solved using offline brute-force search. The optimum solution, denoted by $D(R, P)$, describes the P-R-D behavior of the video encoder. The corresponding optimum complexity control parameters are denoted by $\{\gamma_i^*(R, P)\}$, $1 \leq i \leq L$. Fig. 2 shows the P-R-D functions $D(R, P)$ for two test video sequences, “Foreman” and “Football,” both encoded by our complexity-scalable MPEG-4 encoder at CIF (352×288) size and 10 fps. We used two complexity control parameters, the number of SAD (sum of absolute difference) computations and the fraction of skipped macroblocks (MBs). Here, a fast algorithm, called diamond search, is used for motion estimation [20]. It should be noted that in both plots the encoding power is normalized by the maximum encoder power $P_{\max}$ where no complexity control is applied.

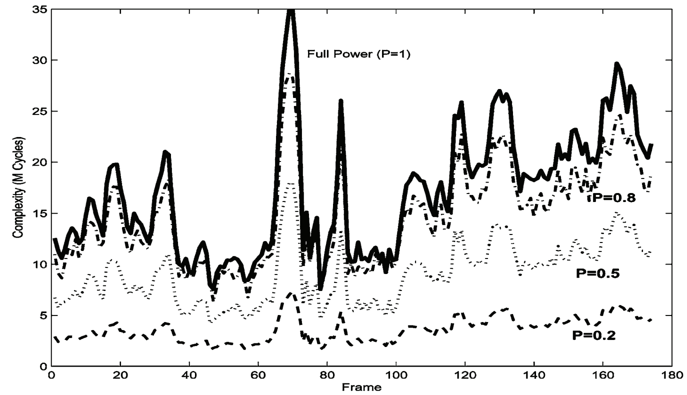


Fig. 3. Encoding complexity of video frames and complexity control.

B. Analytic Power-Rate-Distortion Models

During our extensive operational P-R-D analysis over a wide range of video sequences, we observe that their P-R-D functions have an exponential behavior, as we can see from Fig. 2. In this work, for the convenience of analysis, we propose to approximate them using the following model

$$D(R, P) = \sigma^2 2^{-\lambda R \cdot g(P)}, \quad 0 \leq P \leq 1. \tag{4}$$

Here, σ^2 represents the variance of encoded picture. If it is a motion predicted video frame, σ^2 is the variance of the different picture after motion compensation. λ is a P-R-D model parameter which characterizes the resource (bits and energy) utilization efficiency of the video encoder. $g(P)$ is the inverse power consumption function $\Phi^{-1}(\cdot)$ after proper normalization such that $g(0) = 0$ and $g(1) = 1$. According to (2), we have

$$g(P) = P^{1/\gamma}. \tag{5}$$

For example, in our DeerCam system to be described in Section III-A, $\gamma = 2.5$. For other microprocessors with DVS capabilities, we typically have $1 \leq \gamma \leq 3$ [23].

It should be noted that in this analytic P-R-D model the encoding power P is normalized by the maximum encoder power $P_{\max}$. When P decreases from 1 to 0, the encoder complexity will be scaled down accordingly. In this case, P can be considered as the percentage of maximum encoding power. In this work, we refer to P as the encoder power consumption level. For example, Fig. 3 shows the encoder complexity of a test video sequence being scaled down at four power consumption levels P =1.0, 0.8, 0.5, and 0.2. Here, we cascade the first 60 frames of “Foreman,” “Coastguard,” and “Bike” CIF (352×288) videos to form the test video sequence. We use two parameters, the number of SAD (sum of absolute difference) computations used by diamond motion search and the number of skipped MBs [7] to control the computational complexity of an MPEG-4 encoder. The encoding frame rate is 10 fps. It can be seen that, because of time-varying scene activities, the maximum video encoder complexity (when $P = 1$ or no complexity control is applied) changes dramatically over time. As P decreases from 1 to 0, the encoder complexity of each video frame decreases accordingly.

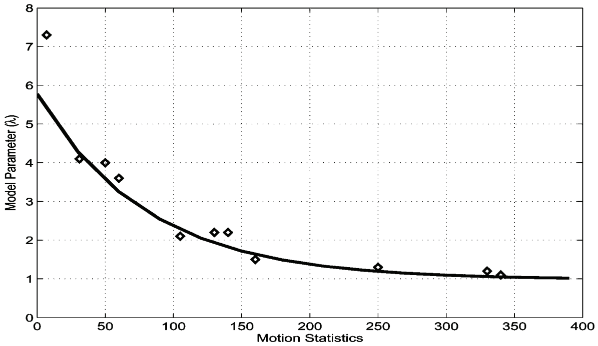


Fig. 4. Correlation between the motion statistics σ_m^2 and P-R-D model parameter λ.

C. Online Estimation of the P-R-D Model Parameter

The P-R-D model in (4) has two major parameters, λ and σ^2. In real-time video encoding, the picture variance σ^2 can be obtained directly from the video encoder. Therefore, we only need to estimate the P-R-D model parameter λ. In Section II-A, we have presented an operational approach to obtain the P-R-D curve. Based on this P-R-D curve, we can determine the value of λ using statistical fitting. This approach is computationally intensive and only suitable for offline P-R-D analysis. In real-time video encoding over portable devices, it is desirable to develop a low-complexity scheme which is able to estimate λ from frame statistics collected at the video encoder.

In this work, we find that the value of λ is highly correlated with motion statistics of video frames. Let $[m_x(i), m_y(i)]$, $1 \leq i \leq M$, be the motion vector of the ith MB in the current video frame. Let $[\bar{m}_x, \bar{m}_y]$ be the mean of all motion vectors. Let ζ be the fraction of Intra MBs in the frame. The motion statistics, denoted by σ_m^2, is defined as

$$\sigma_m^2 = \zeta \cdot 256 + \frac{1}{M}\sum_{i=1}^{M}[m_x(i) - \bar{m}_x]^2 + [m_y(i) - \bar{m}_y]^2. \quad (6)$$

Here, "256" is an empirical number which represents the average motion complexity of an Intra MB. In our experiments, the motion search range is set to be 16 pixels. To study the correlation between the motion statistics σ_m^2 and the P-R-D model parameter λ, we execute the P-R-D video encoder over a wide range of test video segments. These video segments are all extracted from standard CIF video sequences, including "Foreman," "Flowergarden," "Carphone," "Football," "Bike," and "Coastguard". Each video segment has 60 frames with the first frame as an Intra frame and the rest as P frames. For each video segment, we collect its P-R-D data using the operational approach explained in Section II-A, fit the data with the analytic P-R-D model in (4), determine the value of λ, and collect its motion statistics σ_m^2. Fig. 4 shows that the P-R-D model parameter λ is highly correlated with the motion statistics σ_m^2 (in diamonds). The following model is then used to estimate λ:

$$\lambda = C_0 + C_1 \times 2^{C_2\sigma_m^2} \quad (7)$$

where $C_0 = 0.98$, $C_1 = 4.8$ and $C_2 = -0.018$. The model fitting result is also shown in Fig. 4 in a solid line. In practice, the values of these three coefficients can be adjusted based on previous P-R-D statistics. More specifically, for those encoded video frames or segments, we know their coding distortion D, coding bit rate R, the corresponding power consumption P, as well as their motion statistics σ_m^2. Based on these statistics, we can determine the values of λ and update the coefficients C_0, C_1, and C_2 using (7) with statistical fitting.

III. Encoding Energy Minimization on Portable Video Devices

In this section, we use one example to discuss portable video communication system design, understand its energy consumption behavior, and study how the P-R-D analysis presented in the previous section can be applied to minimize video encoding energy consumption.

A. Portable Video Encoding System Design and Energy Characterization

Fig. 5(b) shows an embedded wireless video communication system, called *DeerCam*, developed at University of Missouri, Columbia, for wildlife activity monitoring and behavior analysis [26], [28]. The system is based on a Crossbow Stargate platform [27]. As illustrated in Fig. 5(a), it has the following major components: a low-power USB camera for video capture, an embedded 400 MHz Intel PXA255 XScale microprocessor, a PCMCIA card for 802.11-based wireless data transmission, and a Compact Flash (CF) card for storing/buffering compressed video bit streams. In our project, we deploy the DeerCam system with rugged housing as shown in Fig. 5(b) on animals (deer) for several weeks to collect video samples about their daily activities for food selection analysis, behavior modeling, and interaction tracking. The video data captured by the digital camera is compressed by an MPEG-4 video encoder running on the embedded microprocessor. The compressed bit streams are temporarily stored in the 4.0 GBytes CF card. We deploy a wireless router near a place (e.g., a food station) that the animal visits often. When the animal comes back, for example, within 10–20 meters from the router, the wireless transmission module will be activated to upload the compressed video data to the router and release the storage space in the CF card, as illustrated in Fig. 5(c).

To characterize the energy consumption of major components of the DeerCam system, we use the power measurement setup depicted in Fig. 6(a) and follow a power consumption measurement procedure described in [30]. The operating voltage of the Stargate system is 3.8–5.0 V [27]. During the video capturing and encoding process, only these required device components, such as camera and CF card, are turned on, while other idle components on the Stargate, such as Ethernet connection, are configured to shut off. The video encoding pipeline has been optimized such that memory access is reduced as much as possible. The current draw of the low-power USB camera is about 73 mA. According to our experience, for this type of wildlife activity monitoring videos with a size of CIF (352 × 288) at 7–10 fps,

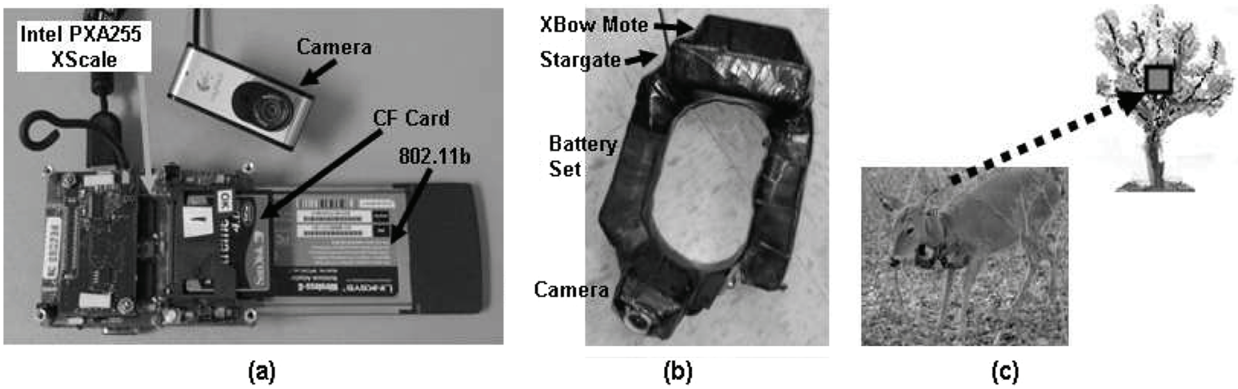


Fig. 5. (a) Major modules of the embedded wireless video communication system. (b) Its final configuration during deployment for wildlife behavior monitoring. (c) Wireless video downloading.

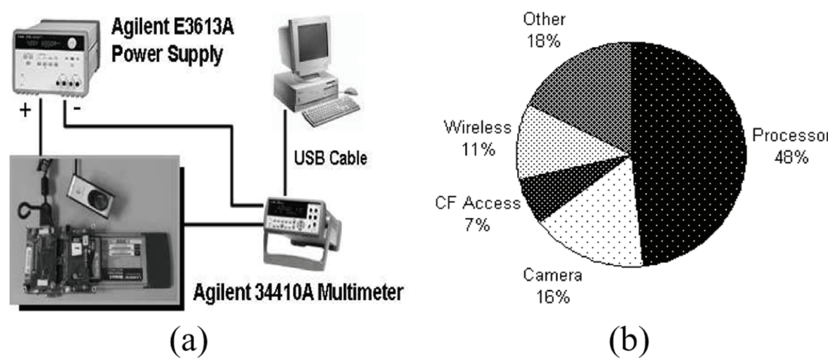


Fig. 6. Average energy consumption of major components of the Stargate system.

the average bit rate is about 400 kbps.[2] During wireless video transmission, the average bandwidth is 2–3 Mbps. Therefore, the required data transmission time is about 10%–20% of the total video encoding time. Fig. 6(b) shows the estimated fractions of energy consumption for major system components. It can be seen that the video encoder (processor) consumes a significant portion of the total energy and wireless transmission energy is about 1/4 of the encoding energy due to delay-tolerant short-range wireless data transmission. Experimental studies in the literature (e.g., [4]) with other communication settings, for example, live video streaming over wireless LAN, also show similar energy consumption behaviors of portable video devices.

B. Energy Tradeoff Between Video Encoding and Wireless Transmission

In this work, we focus on energy minimization of video encoding and wireless data transmission since these two consume a major portion of the total energy while the energy consumption of the remaining components depends on specific system design and is not easily controllable from a video encoding perspective.

According to operational P-R-D analysis in Section II, video encoding power P is a function of encoding bit rate R and distortion D, denoted by $P(R, D)$. The wireless data transmission power P_t is given by

$$P_t = c_t \cdot R \tag{8}$$

where c_t is the energy cost that is needed for successful transmission of one data bit. It depends on transmission distance and path loss index [29], [31]. This is just a simplified model to demonstrate the energy tradeoff between video encoding and wireless transmission. In our case, it is reasonable since the delay is large and the transmission distance is relatively small and we can consider c_t to be the average transmission energy cost. For a given video encoding distortion (or equivalently picture quality) D, if we decrease the encoding power P, the encoder will generate more bits and a higher bit rate R is needed to achieve the target distortion D. According to (8), this implies more power is needed for wireless transmission. As illustrated in Fig. 7, this leads to an energy tradeoff between video encoding and wireless transmission. This suggests that, in many practical video encoding scenarios where the system has access to sufficient storage (or buffer) space or transmission bandwidth, the per-bit energy cost in wireless data transmission is relatively low, while the video encoder operates under severe energy constraints, we can lower down the encoding power to an optimum level O, as illustrated in Fig. 7 with a triangle, to minimize the overall power consumption. This energy tradeoff can conceptually described by the following minimization problem:

$$\min_{R} P(R, D) + c_t \cdot R. \tag{9}$$

For a detailed treatment of this energy tradeoff problem, see [29]. This resource allocation between video encoding and wireless transmission determines the optimum bit budget, denoted by R_T, for video encoding.

[2]This is the average encoding bit rate for video samples of different animal activities, including feeding, bedding, walking, running, etc.

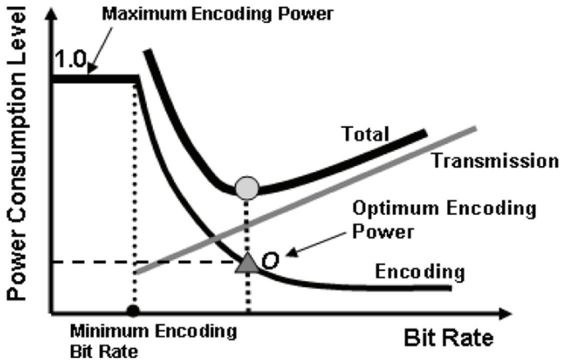


Fig. 7. Energy tradeoff between video encoding and wireless data transmission.

IV. Energy Saving Using P-R-D Optimization

Once the optimum bit budget R_T for video encoding is determined, in this section, we will study how the P-R-D model developed in Section II can be used to minimize the overall encoding energy through optimum allocation of the bit budget R_T within the video sequence. The major result in this section can be summarized as follows: if the input video data is nonstationary with time-varying scene characteristics, the P-R-D model enables us to explore the input source diversity, to trade bits for joules (energy) between different video segments, and to achieve significant encoding energy saving by performing optimum bit allocation.

In practice, a video scene under surveillance or monitoring often experiences a series of events with time-varying scene activity levels. Therefore, the video sequence to be encoded by the portable device is often highly nonstationary. For example, in our DeerCam system, the animal has a wide variety of daily activities, such as feeding, walking, and bedding, which cause significant content changes in the input video. This observation also holds in personal video recording and many remote video surveillance applications. In the following, we will study how this nonstationary nature of video input can be exploited by P-R-D optimization to minimize the energy consumption of video encoding and prolong the operational lifetime of portable video devices.

A. Energy Saving Through Optimum Rate Allocation

We consider two video encoding schemes. In Scheme A, denoted by $\prod_A$, the encoder assumes that the video input is stationary, uses the average P-R-D statistics to determine the optimum encoding power as illustrated in Fig. 7, and encodes all video frames in the session using this constant power, denoted by $\bar{P}$. In Scheme B, denoted by $\prod_B$, the encoder partitions the input video into segments and performs optimum resource allocation among these video segments to minimize the overall encoding power consumption. Both theoretically and experimentally, we are going to demonstrate that, for the same bit budget and video quality, encoder $\prod_B$ requires less energy than $\prod_A$. We will analyze the energy saving gain and discuss how this can be achieved in practical system design. Let S be the video sequence to be encoded. The time duration of S is denoted by T, which corresponds to the operational lifetime of device. We partition S into a number of segments, denoted by $\{S_i | 1 \le n \le N\}$. Suppose that the average picture variance of video segments (after motion compensation) and P-R-D model parameter in (4) for each video segment are σ_i^2 and λ_i, respectively. Let R_i and P_i be the amount of bits and power consumption level used by video encoder $\prod_B$ to compress video segment S_i. Let R_T be the total number of bits that can be generated by the video encoder. We assume that the video sequence is encoded at a constant quality D. In general, this assumption is reasonable because most applications require that videos be encoded at a target quality. Rewriting (4), for video segment S_i, we have

$$P_i = \frac{a_i}{R_i^\gamma} \tag{10}$$

where

$$R_i \ge a_i^{1/\gamma} \tag{11}$$

and

$$a_i = \left(\frac{1}{\lambda_i}\log_2\frac{\sigma_i^2}{D}\right)^\gamma \tag{12}$$

is called *scene activity level*. The constraint in (11) makes sure that $P_i \in [0,1]$. We can see that within the P-R-D analysis framework, the encoder power consumption P_i depends on the encoding bit rate R_i. Our objective is to minimize the total power consumption of encoder $\prod_B$ in compressing all video segments. For the convenience of analysis, let's ignore the rate constraint in (11) at this moment. During online resource allocation in Section V, we will handle this constraint. The energy minimization problem can be formulated as follows:

$$\min_{\{R_i\}} \quad P = \sum_{i=1}^N P_i = \sum_{i=1}^N \frac{a_i}{R_i^\gamma} \tag{13}$$

$$\text{s.t.} \quad \sum_{i=1}^N R_i \le R_T. \tag{14}$$

B. Optimum Encoding Power

To determine the optimum encoding bit rate R_i and power P_i for each video segment S_i, we need to solve the constrained optimization problem in (13) using a Lagrange multiplier approach. Let

$$\mathbf{J} = \sum_{i=1}^N \frac{a_i}{R_i^\gamma} + \beta\left(\sum_{i=1}^N R_i - R_T\right) \tag{15}$$

where β is the Lagrange multiplier. For the minimum solution, the following condition holds

$$\frac{\partial \mathbf{J}}{\partial R_n} = -\gamma\frac{a_n}{R_n^{\gamma+1}} + \beta = 0, \quad 1 \le n \le N. \tag{16}$$

That is

$$R_n = \left(\frac{\gamma a_n}{\beta}\right)^{1/\gamma+1}. \tag{17}$$

Since $\sum_{n=1}^{N} R_n = R_T$, we have

$$\beta^{1/\gamma+1} = \frac{\sum_{n=1}^{N} (\gamma a_n)^{1/\gamma+1}}{R_T}. \tag{18}$$

From (17), we have

$$R_n = \frac{(\gamma a_n)^{1/\gamma+1}}{\sum_{i=1}^{N} (\gamma a_i)^{1/\gamma+1}} R_T \tag{19}$$

which is the optimum bit allocation. The minimum overall power consumption is then given by

$$\begin{aligned} P = \sum_{i=1}^{N} P_n = \sum_{n=1}^{N} \frac{a_n}{R_n^{\gamma}} &= \frac{1}{R_T^{\gamma}} \sum_{n=1}^{N} a_n \frac{\left[\sum_{i=1}^{N} (\gamma a_i)^{1/\gamma+1}\right]^{\gamma}}{(\gamma a_n)^{\gamma/\gamma+1}} \\ &= \frac{1}{R_T^{\gamma}} \left[\sum_{n=1}^{N} a_n^{1/\gamma+1}\right]^{\gamma} \sum_{i=1}^{N} a_i \frac{\gamma^{\gamma/\gamma+1}}{(\gamma a_i)^{\gamma/\gamma+1}} \\ &= \frac{1}{R_T^{\gamma}} \left[\sum_{n=1}^{N} a_n^{1/\gamma+1}\right]^{\gamma} \sum_{i=1}^{N} a_i^{1/\gamma+1} \\ &= \frac{\left[\sum_{i=1}^{N} a_i^{1/\gamma+1}\right]^{\gamma+1}}{R_T^{\gamma}}. \end{aligned} \tag{20}$$

C. Analysis of Energy Saving Gain

In this section, we will demonstrate that, for the same bit budget R_T and video quality D, encoder $\prod_B$ requires less energy than $\prod_A$. As discussed above, in encoder $\prod_A$, a constant encoder power $\bar{P}$ is used. If D is the target video encoding quality, according to (10), the corresponding number of encoding bits of $\prod_A$, denoted by R_i^A, is given by

$$R_i^A = \left(\frac{a_i}{\bar{P}}\right)^{1/\gamma}. \tag{21}$$

Therefore, the total number of bits generated by encoder $\prod_A$ is

$$R_T = \sum_{i=1}^{N} \left(\frac{a_i}{\bar{P}}\right)^{1/\gamma} = \bar{P}^{-1/\gamma} \sum_{i=1}^{N} a_i{}^{1/\gamma} \tag{22}$$

and the total power consumption levels of encoder $\prod_A$ is given by $P_T^A = N \cdot \bar{P}$. For the same number of encoding bit budget R_T and video distortion D, according to (20) and (22), the minimum total power consumption of encoder $\prod_B$ is

$$P_T^B = \frac{\left(\sum_{i=1}^{N} a_i^{1/\gamma+1}\right)^{\gamma+1}}{R_T^{\gamma}} = \bar{P} \cdot \frac{\left(\sum_{i=1}^{N} a_i^{1/\gamma+1}\right)^{\gamma+1}}{\left(\sum_{i=1}^{N} a_i^{1/\gamma}\right)^{\gamma}}. \tag{23}$$

Therefore, the energy saving ratio of encoder $\prod_B$ over $\prod_A$ is then given by

$$\Lambda_e = \frac{P_T^B}{P_T^A} = \frac{\left(\sum_{i=1}^{N} a_i^{1/\gamma+1}\right)^{\gamma+1}}{\left(\sum_{i=1}^{N} a_i^{1/\gamma}\right)^{\gamma} \cdot N}. \tag{24}$$

According to the proof in Appendix A, we have

$$\Lambda_e = \frac{P_T^B}{P_T^A} \le 1 \tag{25}$$

and the equality holds if and only if the scene activity levels $\{a_i\}$ are equal to each other, which implies that the input video is stationary. Note that in the P-R-D model in (4), we have $0 \le P_n \le 1$. However, the above energy minimization has not considered this constraint. It is highly possible that $P_n > 1$ if $\bar{P}$ is very close to 1. If this happens, one option is to crop this optimum value and set it to be $P_n = 1$. In this case, the solution becomes sub-optimum and the energy minimization performance degrades.

V. Online Resource Allocation and Energy Minimization

Section IV provides a theoretical analysis on the energy saving performance of P-R-D optimization. In this section, we discuss how this theoretical energy saving performance can be realized in practical video compression. More specifically, we will discuss how to perform online power allocation and control in real-time video compression. There are three major issues that need to be addressed. 1) First, the optimal resource allocation in Section IV-A requires global knowledge of scene activity levels $\{a_i | 1 \le n \le N\}$ of all video segments, which is infeasible in real-time video communication because we do not have access to future video segments and their statistics. 2) Second, in energy allocation between video encoding and wireless transmission as discussed in Section III-B, we assume that the input video is stationary and its P-R-D behavior does not change over time. 3) Third, once the optimum power consumption level of the video segment is determine, we need to develop a scheme to configure the complexity control parameters to achieve the target power consumption level. In the following, we will address these issues.

A. Predicting Scene Activity Levels of Future Frames for Resource Control

To address the first issue, we propose to model the scene activity levels using a Markov model. Based on this model and past statistics, we predict the scene activity levels a_i of future video segments. More specifically, using the method presented in Section II-C, we are able to estimate the value of λ. According to its definition in (12), we can obtain the value of scene activity level a_n once video segment S_n is encoded. We quantize the value of scene activity level into K discrete levels. Here, for the ease of implementation, a uniform quantizer is used. We then use an K-state Markov model, which has K scene activity states, denoted by $\{A_m | 1 \le m \le K\}$, to model the temporal change of scene activity levels of the video sequence.

Let $[p_{mk}]_{1\le m,k\le K}$ be the state transition matrix and p_m the probability of state A_m. In the proposed online resource allocation scheme, we assume that temporal changing patterns of scene activity levels of future video segments would be similar to those of previous encoded video segments. In this way, the probability distribution of scene activity levels of future video segments would be the same as that of past video segments.

B. Online Power Allocation Between Video Encoding and Wireless Transmission

In the following, we propose an online scheme to explore the tradeoff between video encoding and wireless transmission so as to determine the optimum encoding bit rate target R_T. Suppose we are encoding video segment S_n. Let $\{p_m\}$ be the distribution of scene activity states of previously encoded video segments. According to the P-R-D model in (10) and the scene activity prediction in Section V-A, the average P-R-D behavior of previous video segments is given by

$$\bar{P}(R,D)=\sum_{m=1}^{K} p_m\cdot\frac{A_m}{R^{\gamma}}. \tag{26}$$

Let c_t be the average per-bit energy cost during past wireless transmission sessions. As discussed in Section III-B, the optimum encoding bit rate R_T can be determined by solving the following optimization problem

$$\begin{aligned} R_T &= \arg\min_R \bar{P}(R,D)+P_t \\ &= \arg\min_R\ c_t\cdot R+\sum_{m=1}^{K} p_m\cdot\frac{A_m}{R^{\gamma}}. \end{aligned} \tag{27}$$

This 1-D minimization can be numerically solved using gradient search. Once the target bit rate R_T is determined, using the rate allocation scheme for energy minimization scheme presented in Section IV-A, we can determine the encoding bit rate target for the current video segment. Suppose S_n is the current video segment. Let $\bar{R}_{n-1}$ be the average bit rate used by previously encoded video segments. Note that the total bit budget is $N\cdot R_T$. The following formula, which is adapted from (19), is used to determine the bit budget for video segment S_n

$$R_n=\frac{(\gamma a_n)^{1/\gamma+1}}{\sum\limits_{m=1}^{K} p_m(\gamma A_m)^{1/\gamma+1}}\cdot\frac{N\cdot R_T-(n-1)R_{n-1}}{N-n+1}. \tag{28}$$

C. Configuration of Encoder Complexity Control Parameters

Using (28), we can determine the target encoding bit rate R_n of the video segment to be encoded. Suppose the target video distortion (quality) is D. Using the P-R-D model (10), we can determine the target encoder power consumption level P_n. The next step is how to configure the complexity control parameters of the video encoder to achieve the P-R-D triplet $[P_n,R_n,D]$. During operational P-R-D analysis in Section II-A, we are able to determine the optimum encoder complexity control parameters $\{\gamma_i^*(R,P)\}$ using offline measurements and numerical optimization. However, this approach cannot be applied in real-time video encoding over portable device because it is too computationally intensive. To address this issue, we propose a training-classification based approach. More specifically, Fig. 4 suggests that the P-R-D behaviors of video segments are highly correlated with their motion statistics. This implies that video segments of similar motion statistics will have similar P-R-D behaviors.

During the training stage, we collect a set of training video segments with a wide range of motion statistics. We partition them into a number of clusters according to their motion statistics σ_m^2. According to our experience, 5 to 7 clusters will be sufficient. For each cluster of video segments, we find their average P-R-D function and optimum encoder complexity control parameters using the operational P-R-D analysis. These optimum encoder complexity control parameters for all clusters are then stored in a database. During real-time video encoding, we compute the motion statistics σ_m^2 of the video segment, determine its cluster based on the value of σ_m^2, and then use the average optimum encoder complexity control parameters of that cluster retrieved from the database to control the video encoder.

D. Summary of the Algorithm

The proposed online resource allocation and energy minimization algorithm is summarized in this section. In this work, we assume that the target video distortion (or quality) of the portable video device is D, which is often pre-configured by users. Let c_t be the average per-bit energy cost of wireless data transmission. In the initialization stage, we assume that the distribution of scene activity levels is uniform, which will be updated later on using scene activity statistics obtained from the video encoder. The following steps are performed during online resource allocation, control, and energy minimization:

1) *Collect motion statistics and estimate the P-R-D parameter*. Compute the average motion statistics σ_m^2 of the current video segment. Using **(7)**, estimate the P-R-D parameter λ. This is actually a chicken-egg problem: we cannot collect motion statistics σ_m^2 before a video segment is encoded, however, σ_m^2 is needed for estimating the value of P-R-D parameter and for determining the optimum encoding bit rate and power. To solve this problem, we can use motion statistics of encoded frames in the current video segment and progressively update the value of σ_m^2 and λ as more and more frames are encoded within the segment.
2) *Scene classification and resource allocation*. Using (12), determine the scene activity level a_n of the current video segment and its Markov state. Using (28), compute the bit rate target R_n for the current segment. With (10), determine the encoding power P_n.
3) *Encoder complexity control*. Based on the motion statistics σ_m^2, determine the P-R-D cluster of the current video segment, use the optimum complexity control parameters of this cluster to control the video encoder, as explained in Section V-C.
4) *Model parameters update*. After the current video segment is encoded, based on the statistics of current and previous video segments, update the coefficients C_0, C_1, and C_2 in model (7). Also, update the probability distribution of scene activity states $\{p_m\}$.
5) Repeat the above steps until all video segment are encoded.

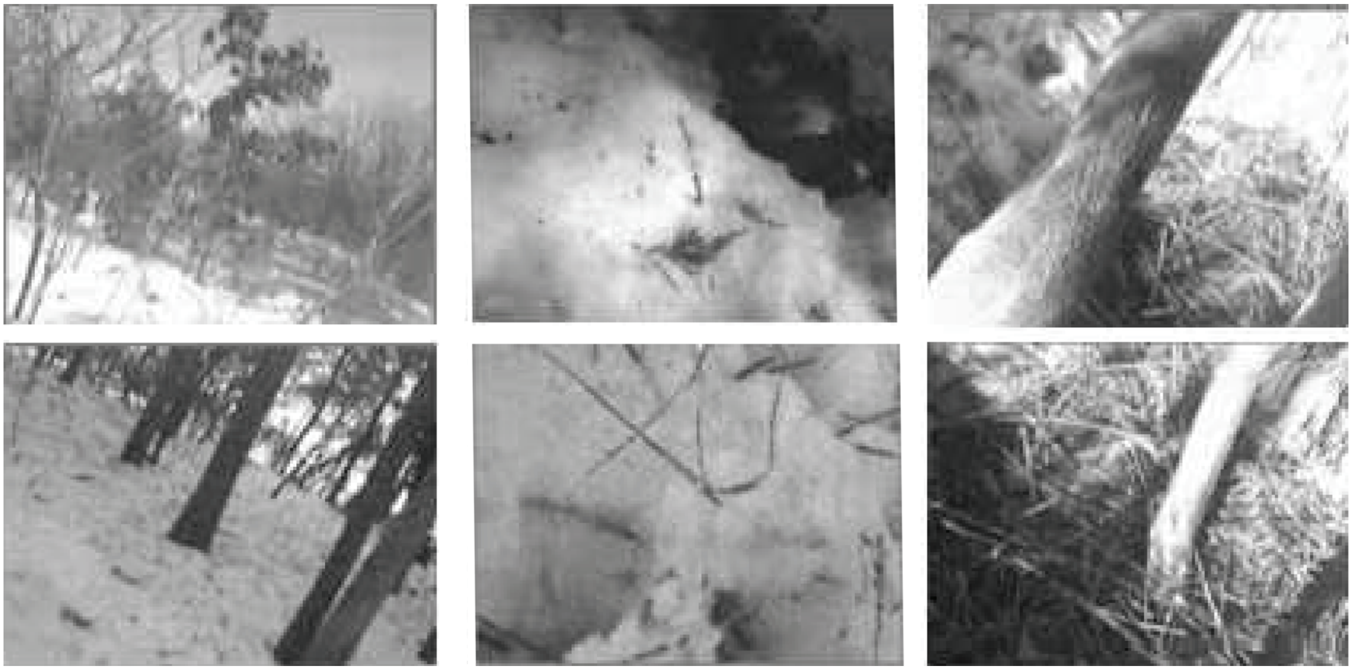


Fig. 8. Sample video frames from the test video sequence captured by the DeerCam system during our field evaluation.

The proposed online P-R-D parameter estimation and energy minimization algorithm operates at the video segment or group of pictures (GOP) level. We recognize that portable video communication device is often expected to operate for hours. Therefore, there is no need to adjust the power allocation at very fine scales, such as frame-level. Instead, it would be sufficient to perform power allocation and control at the video segment level. From a computational complexity perspective, the computation overhead of this type of segment-level resource allocation and energy minimization scheme is very small.

VI. Experimental Results

In this section, we experimentally evaluate the proposed scheme for resource allocation and energy minimization. We will also analyze the sensitivity of energy saving gain to the power consumption model of computing microprocessors. We then study how the wireless transmission energy cost affects the overall energy saving gain.

We have implemented the energy scalable MPEG-4 video encoder described in Section II on a PC. For each video segment, we are able to measure the number of cycles used by the video encoder. At this moment, we have not ported this energy-scalable encoder onto the Stargate platform and linked it with its online dynamic voltage control. Instead, we use the power consumption model depicted in Fig. 1, which is obtained during our offline study, to map the computational complexity into encoder power consumption. The test videos used in this experiment are deer activity monitoring videos. In November 2006, we deployed the DeerCam system with a nonenergy-scalable MPEG-4 encoder on deer and obtained more than 60 h of animal activity videos of CIF size at approximately 5 fps. Here, the frame rate is relatively low, however, it is sufficient for wildlife activity analysis. In addition, a low frame rate requires less encoding power and enables us to monitor animals over an extended period of time. In Table I, we list 6 of these test video sequences which are labeled by its day and time, as well as the total number of frames we have tested. For example, the test video sequence "DeerCam D1(AM)" means that this video was captured at day 1 in the morning. We used the first 40 000 frames from this sequence during our experiment. Fig. 8 shows a sample video frame for each of these DeerCam videos. Besides these 6 DeerCam videos, we also include a standard test video sequence, labeled as "MPEG," which is generated by cascading the first 150 frames of 6 CIF videos: "Foreman," "Carphone," "Flowergarden," "Coastguard," "NBA," and "News." During our experiments, each video segment has 60 frames with the first frame as an I-frame and the rest as P-frames.

TABLE I
PERFORMANCE EVALUATION OF ENERGY MINIMIZATION

Video Sequence	Number of Frames	Relative Rate Control Error	Energy Saving Ratio
MPEG	900	-3.2%	0.61
DeerCam D1(AM)	40000	1.2%	0.56
DeerCam D1(Noon)	25000	-2.3%	0.62
DeerCam D1(PM)	25000	1.8%	0.59
DeerCam D2(AM)	35000	3.2%	0.68
DeerCam D2(Noon)	25000	3.3%	0.53
DeerCam D2(PM)	25000	-2.8%	0.69

A. Performance Evaluation of Resource Control and Energy Minimization

First, we evaluate the energy saving performance of the proposed scheme. Fig. 9 (top) shows the motion statistics σ_m^2 of each video frame of "DeerCam D1(AM)."Fig. 9 (bottom) shows the motion statistics from frames 20 000 to 20 250. It can be seen that the video scene activity is highly nonstationary and changes dramatically over time. To evaluate the performance of P-R-D parameter estimation as described in Section II-C, which is one of the key components of our proposed algorithm,

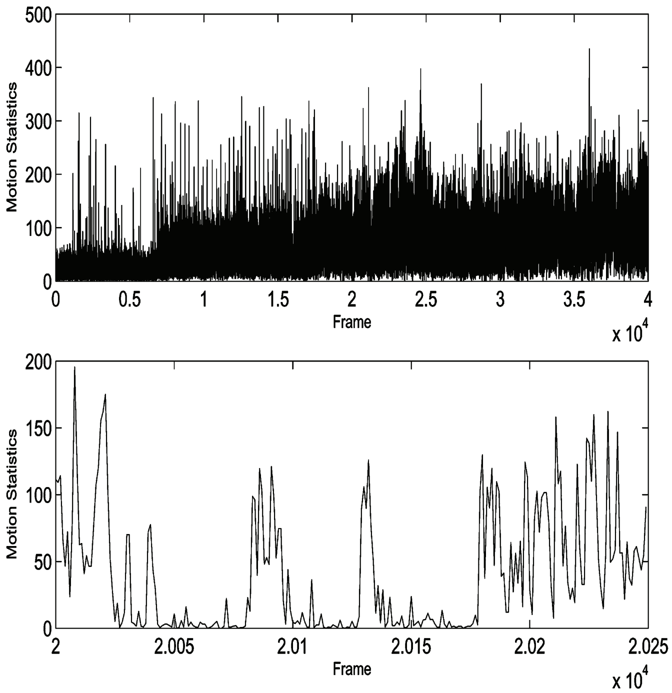


Fig. 9. Motion statistics σ_m^2 of each video frame.

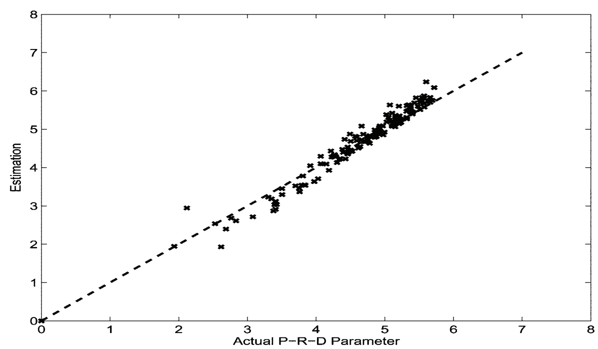


Fig. 10. Estimation results of the P-R-D parameter λ.

we perform the operational P-R-D analysis over 90 video segments randomly selected from the test videos, and obtain the true values of their P-R-D parameter λ. We then compare them with those estimated by the formula in (7). Fig. 10 shows that the estimated values of λ are very close to their actual values.

We set the target video quality to be 33 dB, i.e., $D = 32.5$ (mean squared error). Fig. 11 shows the encoding bit rate (in megabits per second) of each video segment. The top plot shows the offline bit allocation results given by (19) while the bottom one shows the actual bit rate used by each video segment controlled by our online resource allocation and control algorithm in Section V-D. Fig. 12 shows the results for offline and online power allocation. It can be seen that in the offline case, the normalized encoding power sometimes is larger than 1.0 which is invalid. However, in the online case, the maximum power consumption level constraint is satisfied. We compare the proposed algorithm against encoder $\prod_A$ (described in Section IV-A) where a fixed encoding power is used without bit allocation for energy minimization. Table I shows the overall energy saving ratio for each test video sequence, as well as

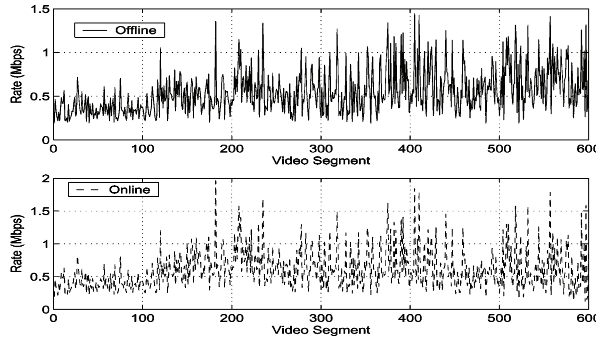


Fig. 11. Bit allocation of each vide segment.

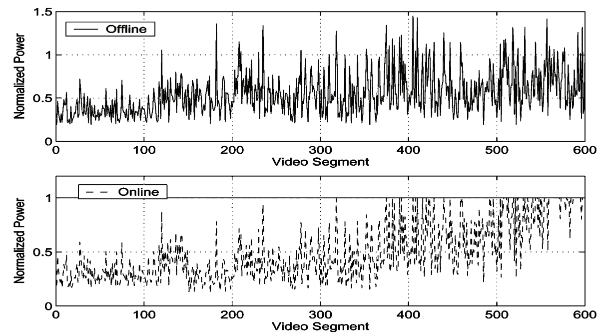


Fig. 12. Power allocation of each video segment.

the rate control results. We can see that for the same video quality and bit budget, the proposed algorithm is able to save about 30%–50% of encoding energy, which is quite significant. It should be noted that this encoding energy saving is made possible by the proposed resource allocation and P-R-D optimization scheme. It is achieved by exploring the nonstationary nature of input video by trading bits for joules between video segments, therefore, is independent of existing software and/or hardware techniques for video encoder optimization. In other words, it achieves further significant encoding energy saving (about 30%–50%) on top of existing software and hardware energy minimization schemes.

B. Sensitivity Analysis of Energy Saving Gain

In the following, we analyze the sensitivity of energy saving ratio to different microprocessor power consumption models, i.e., different values of γ. In the above experiments, the value of γ is 2.5. Fig. 13 shows the histogram of scene activity levels $\{a_i\}$ of all video segments in "DeerCam D1(AM)". It can be seen that it follows a Gamma distribution. Based on this observation, for the convenience of sensitivity analysis purpose, we approximate the distribution of scene activity levels $\{a_i\}$ using the following Gamma distribution:

$$\text{Prob}\{\text{Scene Activity level} = x\} = p(x; a, \theta) = x^{a-1}\frac{e^{-x/\theta}}{\theta^a \Gamma(a)} \tag{29}$$

where a is the shape parameter and θ is the scale parameter of the Gamma distribution. Its mean and variance are given by $a\theta$ and $a\theta^2$, respectively. Based on this distribution, we generate a large number of Gamma random processes, each representing

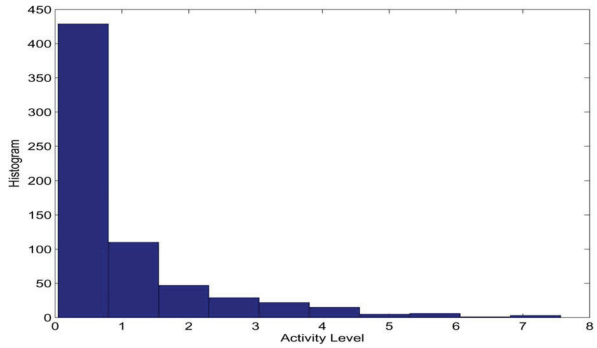


Fig. 13. Distribution of activity levels of video segments.

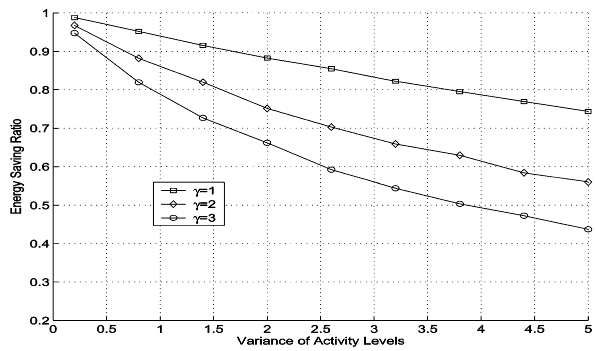


Fig. 14. Energy saving ratios of the P-R-D video encoder at different distributions of scene activity parameters.

the scene activity levels $\{a_n\}$ of a simulated video input. For each simulated video input, we compute the energy saving ratio using (24) with different values of microprocessor power consumption model parameters γ. In Fig. 14, we plot the energy saving ratio at different variance of scene activity levels and different values of γ. Here, we set the average scene activity level to be 2. It can be seen that when the variance of scene activity level is larger, which implies that the input video has a larger variation in its scene content, the energy saving is more significant. As we can see from the proof in Appendix A, the energy saving ratio given by (24), becomes 1 when all scene activity levels are equal to each other. In this case, their variance is zero. When their variance becomes larger, (24) moves further aways from this extreme condition and the energy saving ratio becomes smaller than 1 (more energy saving). We can also see that when γ is larger, which implies that the power consumption of the microprocessor is more sensitive to its work load or application complexity, the overall energy saving is more significant.

VII. Concluding Remarks

In this work, based on P-R-D optimization, we develop a new approach for energy minimization for delay-tolerant video communication over portable devices. Theoretically, we demonstrated that extending the traditional R-D analysis to P-R-D analysis gives us another dimension of flexibility in resource allocation and performance optimization. Within the new P-R-D optimization framework, the bit and energy resources are tightly coupled, which enables us to trade "bits" for "joules" (energy) so as to save energy in data compression. We have analyzed the energy saving performance of P-R-D optimization. We have also developed an adaptive scheme to estimate the P-R-D model parameters and perform online energy optimization and control for real-time video compression. Our simulation results show that, for typical videos with nonstationary scene statistics, using the proposed P-R-D optimization technology, the video encoding energy can be significantly reduced, especially for delay-tolerant video communication applications where the per-bit energy cost of wireless transmission is relatively low. This has a significant impact on energy-efficient portable video communication system design.

In our future work, we shall study and evaluate resource allocation and energy minimization for real-time video encoding and communication on embedded computing platforms with dynamic voltage control. In this work, we focus on application-layer encoder control and optimization. In our next step of research, we shall also study how the P-R-D analysis could be used for joint hardware-layer scheduling and application-layer encoder control to minimize the overall system energy consumption.

Appendix

We prove that the energy saving ratio Λ_e in (24) is less than or equal to 1. To this end, we need to use the Holder's inequality: Given two N-dimensional vectors $(x_1, x_2, \ldots, x_i)$ and $(y_1, y_2, \ldots, y_N)$, according to Holder's inequality, we have

$$\sum_{i=1}^{N} |x_i \cdot y_i| \leq \left(\sum_{i=1}^{N} x_i^p\right)^{1/p} \cdot \left(\sum_{i=1}^{N} y_i^q\right)^{1/q} \qquad (30)$$

where $p, q > 1$ and

$$\frac{1}{p} + \frac{1}{q} = 1. \qquad (31)$$

The equality holds when

$$(x_1, x_2, \ldots, x_N) = \alpha \cdot (y_1, y_2, \ldots, y_N) \qquad (32)$$

where α is a constant. We let

$$x_i = a_i^{1/\gamma+1}, \quad y_i = 1 \qquad (33)$$

$$p = \frac{\gamma+1}{\gamma}, \quad q = \gamma + 1. \qquad (34)$$

Using the Holder's inequality, we have

$$\sum_{i=1}^{N} \left| a_i^{1/\gamma+1} \cdot 1 \right| \leq \left[\sum_{i=1}^{N} \left(a_i^{1/\gamma+1}\right)^{\gamma+1/\gamma}\right]^{\gamma/\gamma+1} \cdot \left(\sum_{i=1}^{N} (1)^{\gamma+1}\right)^{1/\gamma+1} \qquad (35)$$

$$= \left[\sum_{i=1}^{N} a_i{}^{1/\gamma}\right]^{\gamma/\gamma+1} \cdot N^{1/\gamma+1}. \qquad (36)$$

That is

$$\left(\sum_{i=1}^{N} a_i^{1/\gamma+1}\right)^{\gamma+1} \leq \left[\sum_{i=1}^{N} a_i{}^{1/\gamma}\right]^{\gamma} \cdot N. \tag{37}$$

Therefore

$$\Lambda_e = \frac{\left(\sum\limits_{i=1}^{N} a_i^{1/\gamma+1}\right)^{\gamma+1}}{\left(\sum\limits_{i=1}^{N} a_i^{1/\gamma}\right)^{\gamma} \cdot N} \leq 1. \tag{38}$$

According to (32), the equality holds when

$$(a_1, a_2, \ldots, a_N) = \alpha \cdot (1, 1, \ldots, 1) \tag{39}$$

which implies that the scene activity level a_i of each video segment is constant.

REFERENCES

[1] (2007). [Online]. Available: http://www.abiresearch.com/products/market_research/Mobile_Broadcast_Video

[2] in *Panel Report of NSF Workshop on Sensors for Environmental Observatories*, Seattle, WA, Dec. 2, 2004 [Online]. Available: http://www.wtec.org/seo

[3] W. Yuan, K. Nahrstedt, S. V. Adve, D. L. Jones, and R. H. Kravets, “Grace-1: Cross-layer adaptation for multimedia quality and battery energy,” *IEEE Trans. Mobile Comput.*, vol. 50, no. 7, pp. 799–815, Jul. 2006.

[4] X. Lu, Y. Wang, and E. Erkip, “Power efficient H.263 video transmission over wireless channels,” in *Proc. ICIP*, Rochester, NY, Sep. 2002, pp. 533–536.

[5] T. Sikora, “The MPEG-4 video standard verification model,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 7, no. 1, pp. 19–31, Feb. 1997.

[6] T. Wiegand, *Text of committee draft of joint video specification* ITU-T rec. H.264—ISO/IEC 14496-10 Avc, Doc. JVTC167, 3rd JVT Meeting, May 2002.

[7] Z. He, Y. Liang, L. Chen, I. Ahmad, and D. Wu, “Power-rate-distortion analysis for wireless video communication under energy constraint,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 15, no. 5, pp. 645–658, May 2005.

[8] P. Agrawal, J.-C. Chen, S. Kishore, P. Ramanathan, and K. Sivalingam, “Battery power sensitive video processing in wireless networks,” in *Proc. IEEE PIMRC*, Boston, MA, Sep. 1998, pp. 116–120.

[9] H. Kim, N. Kamaci, and Y. Altunbasak, “Low-complexity rate-distortion optimal macroblock mode selection and motion estimation for MPEG-like video coders,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 15, pp. 823–834, Jul. 2005.

[10] D. S. Turaga, M. van der Schaar, and B. Pesquet-Popescu, “Complexity scalable motion compensated wavelet video encoding,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 15, no. 8, pp. 982–993, Aug. 2005.

[11] I. M. Pao and M. T. Sun, “Statistical computation of discrete cosine transform in video encoders,” *J. Vis. Commun. Image Represent.*, vol. 9, no. 2, pp. 163–170, Jun. 1998.

[12] J. Villasenor, C. Jones, and B. Schoner, “Video communications using rapidly reconfigurable hardware,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 5, no. 6, pp. 565–567, Dec. 1995.

[13] W. P. Burleson, P. Jain, and S. Venkatraman, “Dynamically parameterized architecture for power-aware video coding: Motion estimation and dct,” in *Proc. 2nd USF Int. Workshop Digital Computat. Video*, 2001, pp. 4–12.

[14] D. G. Sachs, S. Adve, and D. L. Jones, “Cross-layer adaptive video coding to reduce energy on general-purpose processors,” in *Proc. ICIP*, Barcelona, Spain, Sep. 2003, pp. 109–112.

[15] V. Akella, M. van der Schaar, and W.-F. Kao, “Proactive energy optimization algorithms for wavelet-based video codecs on power-aware processors,” in *Proc. IEEE ICME*, Jul. 2005, pp. 566–569.

[16] A. N. Skodras, “Fast discrete cosine transform pruning,” *IEEE Trans. Signal Process.*, vol. 42, no. 7, pp. 1833–1837, Jul. 1994.

[17] K. R. Rao and P. Yip, *Discrete Cosine Transform—Algorithms, Advantages, Applications*. New York: Academic, 1990.

[18] F. Pan, X. Lin, R. Susanto, K. P. Lim, Z. G. Li, G. N. Feng, D. J. Wu, and S. Wu, “Fast mode decision algorithm for intraprediction in H.264/avc video coding,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 15, no. 7, pp. 813–822, Jul. 2005.

[19] C. Grecos and M. Yang, “Fast mode prediction for the baseline and main profiles in the H.264 video coding standard,” *IEEE Trans. Multimedia*, vol. 8, pp. 1125–1134, Dec. 2006.

[20] B. Zeng, Li Renxiang, and M. L. Liou, “Optimization of fast block motion estimation algorithms,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 7, no. 8, pp. 833–844, Dec. 1997.

[21] P.-C. Tseng, Y.-C. Chang, Y.-W. Huang, H.-C. Fang, C.-T. Huang, and L.-G. Chen, “Advances in hardware architectures for image and video coding—A survey,” *Proc. IEEE*, vol. 93, no. 1, pp. 184–197, Jan. 2005.

[22] M. v. d. Schaar and Y. Andreopoulos, “Rate-distortion-complexity modeling for network and receiver aware adaptation,” *IEEE Trans. Multimedia*, vol. 7, no. 6, pp. 471–479, Jun. 2005.

[23] T. Burd and R. Broderson, “Processor design for portable systems,” *J. VLSI Signal Process.*, vol. 13, no. 2, pp. 203–222, Aug. 1996.

[24] P. Pirsch and H.-J. Stolberg, “VLSI implementations of image and video multimedia processing systems,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 8, no. 5, pp. 878–891, Nov. 1998.

[25] Z.-L. He, C.-Y. Tsui, K.-K. Chan, and M. Lion, “Low-power VLSI design for motion estimation using adaptive pixel truncation,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 10, no. 5, pp. 669–678, Aug. 2000.

[26] University of Missouri, Columbia (2008). [Online]. Available: http://videonet.ece.missouri.edu/deernet/index.htm

[27] Crossbow Stargate Platform (2006). [Online]. Available: http://www.xbow.com/Products/Product_pdf_files/Wireless _pdf/Stargate _Datasheet.pdf

[28] R. J. Moll, J. Millspaugh, J. Beringer, J. Sartwell, and H. Zhihai, “Animal-borne video systems: a new era of behavioral ecology,” *Trends Ecol. Evol.*, vol. 22, no. 12, pp. 660–668, 2007.

[29] Z. He and D. Wu, “Resource allocation and performance limit analysis of wireless video sensors,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 16, no. 5, pp. 590–599, May 2006.

[30] C. B. Margi, V. Petkov, K. Obraczka, and R. Manduchi, “Characterizing energy consumption in a visual sensor network testbed,” in *Proc. 2nd Int. Conf. TRIDENTCOM*, Mar. 2006, pp. 8–15.

[31] T. S. Rappaport, *Wireless Communications: Principles and Practice*. Englewood Cliffs, NJ: Prentice Hall, 1996.

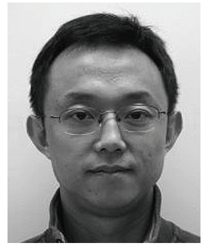
Zhihai He (SM’06) received the B.S. degree from Beijing Normal University, Beijing, China, and the M.S. degree from Institute of Computational Mathematics, Chinese Academy of Sciences, Beijing, China, in 1994 and 1997, respectively, both in mathematics, and the Ph.D. degree from University of California, Santa Barbara, CA, in 2001, in electrical engineering.

In 2001, he joined Sarnoff Corporation, Princeton, NJ, as a Member of Technical Staff. In 2003, he joined the Department of Electrical and Computer Engineering, University of Missouri, Columbia, as an assistant professor. His current research interests include image/video processing and compression, network transmission, wireless communication, computer vision analysis, sensor network, and embedded system design.

Dr. He received the 2002 IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY Best Paper Award, and the SPIE VCIP Young Investigator Award in 2004. He is a member of the Visual Signal Processing and Communication Technical Committee of the IEEE Circuits and Systems Society, and serves as Technical Program Committee member or session chair of several international conferences.

Wenye Cheng received the B.S. degree in electrical engineering from Harbin Institute of Technology (HIT), Harbin, China, in 1998, the M.E. degree from Aerospace institution of China, China, in 2001, and the M.S. degree in electrical and computer engineering from the New Jersey institute of Technology, Newark, in 2004. He is currently working towards the Ph.D. degree in the Department of Electrical and Computer Engineering, University of Missouri, Columbia.

His main research interests include video compression/coding, video surveillance and intelligent video processing. Currently, he is working on the optimized power-rate-distortion model for video encoder.



Xi Chen received the B.S. degree in infrared technologies from Xidian University, Xi'an, China, in 1983, and the M.S. degree from the Electrical and Computer Engineering of University of Missouri, Columbia, in 2006.

He is currently with Harmonic Inc. , Sunnyvale. His research interests include embedded system design for image/video compression, processing, and communication.