Energy-Aware Portable Video Communication System Design for Wildlife Activity Monitoring

Zhihai He, Jay Eggert, Wenye Cheng,
Xiwen Zhao, Joshua Millspaugh,
Remington Moll, Jeff Beringer, and Joel Sartwell

Abstract

In this paper, we introduce our recent research and development effort on energy-efficient portable video communication system design for wildlife activity monitoring. The capability of seeing what an animal sees in the field is very important for wildlife activity monitoring and research. We design an integrated video and sensor system, called *DeerCam* and mount it on animals so as to collect important video and sensor data about their activities in the field. From the video and sensor data collected by *DeerCam*, wildlife researchers will be able to extract a wealth of scientific data for studying the behavior patterns of wildlife species and understanding the dynamic of wildlife systems. We present the system architecture of *DeerCam*, explain our system design goals, and discuss major design issues. One of the central challenges in *DeerCam* system design is energy minimization. We present a new approach for energy minimization of portable video devices: power-rate-distortion (P-R-D) analysis and optimization. We discuss various approaches to minimizing the energy consumption of *DeerCam*, which can be also applied to other portable video devices. Results demonstrate that, by incorporating the third dimension of power consumption into conventional rate-distortion (R-D) analysis, P-R-D analysis gives us one extra dimension of flexibility in resource allocation and energy minimization and allows us to significantly reduce energy consumption.

Key Words: video coding, energy minimization, embedded system, environmental tracking.

Digital Object Identifier 10.1109/MCAS.2008.923007

I. Introduction

The biological relationship between wildlife and humans has never been more intertwined. Outbreaks of infectious wildlife diseases, such as chronic wasting disease (CWD), threaten wildlife populations, human life, food safety and our national economy [1]. Due to the limitations of current wildlife monitoring technologies, the behavior of free-ranging animals, the dynamics of wildlife systems, and the spread of some wildlife diseases remains largely unknown to us. Lack of scientific knowledge about the behavioral interactions and dynamics of wildlife systems significantly limits our ability to effectively manage wildlife resources and control wildlife diseases. For example, in current practice, one choice to fight against wildlife diseases such as CWD is removal of all individuals within an area of an outbreak, since we have little knowledge about how the disease might propagate within the dynamic wildlife system. Clearly, this approach is ineffective and damages the national economy. Therefore, there is a pressing need to develop a new generation of technologies to monitor wildlife behaviors and interactions and study the dynamics of wildlife systems.

In the past few decades, engineers and wildlife researchers have been developing advanced communication technologies for wildlife monitoring [1, 2]. The

state-of-the-art technologies for wildlife monitoring are radio tracking and sensor networks. Radio tracking uses a radio transmitter or receiver (attached to an animal) to collect its location information [1]. There are three distinct types of radio-tracking technologies that are in use today: *very high frequency (VHF) radio tracking* [3], *satellite tracking* [4], and *Global Positioning System (GPS) tracking* [1, 5]. Recent technological advances in hardware miniaturization of sensors, low-power micro-processor design, and wireless ad hoc networking have enabled the deployment of large-scale wireless sensor networks (WSN's) [6, 7, 9, 10]. The WSN technology provides the research community an enabling platform to simultaneously monitor a large group of free-ranging animals at granular scales. For example, ZebraNet, developed at Princeton University [5], utilizes GPS-based radio tracking and sensor networks to track the movement of a group of zebras and study animal migrations and inter-species interactions. GPS receivers are also used for animal tracking in Telenor R&D's Electronic Shepherd project [8] and UC Davis's Southern California Puma Project [11]. Embedded sensor networks have been developed by the research teams at UC Berkeley [13], UCLA [12], as well as in Australia [14] for habitat and health monitoring, using sensors to collect temperature, humidity, and other biological information on research animals.

However, the aforementioned wildlife monitoring technologies do not offer visual information. Looking back at the history of wildlife research, we find that wildlife researchers have been pursuing a dream, a dream to *see* what the animals *see* in the field without disturbing their natural behaviors, in a cost-effective way. Wildlife researchers have found that for accurate behavior analysis and interaction modeling, it is imperative to obtain some visual information about the animal's activity, its resource selection, as well as the environmental context of the behavior [15]. Otherwise, we are kept "blind" from the animals and fail to understand important behavioral attributes of wildlife species. For example, with animal's location only,

we do not know what the animal is doing, how it is doing, and why it is doing like this. Recent research [18] also shows that by not considering what the animal is doing at these point locations, the researcher may obtain a biased estimate of animal's resource selection and behavior patterns. Wildlife researchers have also observed that pooling across behaviors within different environmental contexts will de-emphasize the importance of some habitats for critical behaviors [18]. This is perhaps the most important technological challenge in wildlife tracking studies.

Another important reason for collecting visual information about animals' activities is interaction modeling for disease tracking and control. Most *epidemic wildlife diseases* spread through direct animal contacts (e.g., via direct contact with saliva) and animals' interactions with the environment and other species. With current technologies, all we have available for analysis is a series of approximate location and movement estimates for a number of animals, and we do NOT know whether animals are directly interacting or not. By collecting and analyzing direct visual information, we can study the animal's close interactions with other animals and their environment, which enables us to understand the dynamics of wildlife systems and to develop quantitative propagation models for wildlife diseases.

In this work, we will present our recent effort on low-power portable video communication system design for wildlife activity monitoring and interaction tracking. More specifically, in this research project¹, we design an integrated video encoding and sensor data collection system, called *DeerCam*, and deploy it on animals to collect important video and sensor information about animals' daily activities in the field over an extended period of time. With intelligent information analysis and fusion, we are able to study animals' behavior patterns and answer many challenging wildlife research questions. We present our DeerCam system architecture, discuss the major design challenges, and present our approaches to addressing these challenges.

The rest of the paper is organized as follows. Section 2 presents the design goals and major design challenges of our DeerCam system. Section 3 presents the DeerCam system architecture. In Section 4, we explain our major approach to energy-efficient portable video communication system design for wildlife activity monitoring. Section 5 concludes the paper.

¹This project is funded by National Science Foundation in collaboration with University of Florida.

Zhihai He (HeZhi@missouri.edu), Jay Eggert, Wenye Cheng, and Xiwen Zhao are with the Department of Electrical and Computer Engineering, University of Missouri, Columbia, MO, 65211. Josh Millspaugh and Remington Moll are with the Department of Wildlife and Fisheries, University of Missouri, Columbia, MO, 65211. Jeff Beringer and Joel Sartwell are with the Resource Science Center, Missouri Department of Conservation, Columbia, MO, 65211.

II. System Framework, Design Goals and Challenges

In this section, we describe the proposed system framework for wildlife activity monitoring, explain our design goals, and discuss major design challenges.

II-A. System Framework

Figure 1 illustrates our proposed framework for wildlife activity monitoring. We design an integrated video encoding and sensor data collection system, called *DeerCam*, and mount it on animals. The *DeerCam*, which will be explained in more detail in Section III, captures video information about animal's daily activities in the field, compresses the video data and stores the bit stream in an on-board storage device. We realize direct wireless data transmission is very difficult because wireless signals attenuate dramatically when the animal moves away, especially in forested terrain. According to our experience, the average transmission distance is only about 20 meters in a wooded area. Therefore, we choose to temporarily store the compressed video data on-board. We also set up several wireless access points near places where the animal visits often, for example, food stations. When the animal returns repeatedly to one of these access points, the compressed video data is uploaded through wireless transmission to a video server in a ground monitoring unit, which connects to the wireless access point via an Ethernet cable. In addition, at the completion of the planned experiment, the *DeerCam* system can be physically retrieved and the video data can be then directly downloaded to a computer for further analysis.

Besides video data, the *DeerCam* system is also equipped with a suite of sensors, such as accelerometer, GPS, temperature, and light sensors, on-board to collect various aspects of information regarding animal activity in the field. From the video data, assisted by intelligent computer vision algorithms (e.g. objects detection and classification), we will be able to extract important information regarding animal food selection and activity patterns, fuse it with sensor data, and use the resulting data to study animal's behavior patterns at different times, locations, and environmental contexts and answer important wildlife research questions pertaining to wildlife resource management, interaction tracking, and disease propagation modeling.

We call the proposed system for wildlife activity monitoring *DeerCam* and select deer as the study subject for the following three major reasons. First, recent CWD outbreak and associated national initiatives are related to deer [15]. For example, recent outbreaks of CWD in many U.S. states (e.g., Colorado, South Dakota, Illinois, Nebraska, Kansas, Wisconsin, etc.) have caused significant economical loss and posed a potential threat to human health

[25]. Second, deer are the most ubiquitous large mammal in the U.S., thus deer diseases potentially impact everyone in the nation. Third, deer are simply a starting point for this research and eventually the technologies developed in this project will be applied to other species, especially with the advances in hardware miniaturization of sensors [9, 26].

II-B. Design Goals and Challenges

There are two basic approaches to wildlife activity monitoring: *passive* and *animal-mounted*. In the passive approach (also called camera traps), the camera and sensor unit is deployed on trees or other fixtures to capture images of animals when they pass by. This passive approach, because of its relative simplicity in system design and deployment, has been widely used in wildlife activity monitoring. For a detailed review of passive camera traps, see [16, 17]. The animal-mounted approach, which mounts the camera and sensor unit directly on live animals, is much more technically challenging. However, it has the capability of seeing what the animal sees in the field and collecting direct and exact information about animal activities and their environmental contexts. Furthermore, because the system moves with the animal, it is able to observe animal activities anywhere anytime, revealing rare behaviors that we have never been able to see before.

II-B-a. Design Goals

In this project, we choose the second animal-mounted approach because of its significant impact in wildlife activity monitoring. Our design goal is *long-lived* and *unobtrusive* wildlife video monitoring. We would like to collect the animals' activity information under various

environmental settings for six months. The long observation period is typically necessary for wildlife behavior study, especially for behavior pattern analysis across multiple seasons [5, 15]. Unobtrusive observation is essential for studying natural behaviors of animals. To this end, the sensing devices attached to animals should be light-weight such that the impact on animals' health and the disturbance to their natural behaviors are minimized. A generally accepted rule is that the weight of attached devices (including battery) should be less than 5% of the animal's body weight [23]. Smaller percentages are recommended for avifauna due to the increased energetic demands of flight and the sensitivity of avian reproductive success [24]. Another important aspect of unobtrusiveness is that, once the device is deployed, it should not require re-capture of the animal to change batteries or re-configure the system.

II-B-b. Design Challenges

The unique requirements in wildlife activity monitoring pose significant challenges in system design. More specifically, we face the following two major challenges: (1) energy consumption minimization, and (2) intelligent resource allocation and sensor data utility maximization. In the following, we will explain these two challenges in more detail.

(1) Energy consumption minimization. As discussed above, the DeerCam system is expected to operate for an extended period of time, for example 6 months, and more importantly, battery change is not allowed once the system is deployed. According to our experience, batteries typically contribute to a majority (about 70–90%) of system weight, as we can see from Figure 3. The *unobtrusive-*

ness requirement indicates that the system weight should be kept as light as possible. This implies that the number of batteries attached to the system should be kept as few as possible and the amount of energy supply is thus very limited. In video-based wildlife activity monitoring, digital video data is voluminous, which has to be efficiently compressed before being stored or transmitted. However, efficient video compression often involves highly sophisticated motion prediction, transform, and coding schemes and is computationally intensive and energy-consuming [19]. Therefore, energy minimization of video encoding and lower-power system design is the most important challenge in DeerCam system design.

(2) Intelligent resource allocation and utility maximization. The objective of video-based wildlife monitoring is to collect important visual information about animals' activities in the field for behavior modeling and other wildlife research tasks. Therefore, the overall system performance should be measured by the utility of video data collected by the DeerCam system for wildlife research purposes (e.g. behavior modeling). Therefore, our design objective is to maximize the utility function under resource constraints. As we will discuss in more detail in Section IV-C-b, the system resource (bit and energy) consumption is controlled by a set of parameters. How to model the inherent relationship between resource control parameters and the video data utility, and how to optimally configure these control parameters so as to maximize the utility function under resource constraints are challenging issues in DeerCam system design.

III. DeerCam System

In this section, we introduce our DeerCam system design, explain its central mechanism in resource control, and characterize its behavior in resource utilization.

III-A. DeerCam System Design

Figure 2 shows the camera and sensor unit of our DeerCam system. The size of this unit is about the size of a PDA. It is based on Crossbow Stargate and Mote [27]. It has the following major components: an embedded 400 MHz Intel PXA255 XScale microprocessor, a USB-camera for video capture, a PCMCIA card for 802.11-based wireless data transmission, a Compact Flash (CF) card for storing compressed video bit streams, and a multi-sensor board with accelerometer, GPS, temperature, and light sensors. The system also includes a drop-off control unit whose functionality will be explained in the following section. The video data captured by the digital camera is compressed by an energy-efficient MPEG-4 video encoder [19] operating on the embedded microprocessor. The compressed bit streams are temporarily stored in the

Figure 2. Major components in the camera and sensor unit of the DeerCam.

4.0 G Bytes CF card. We deploy several wireless routers near places (e.g. a food station) that the animal visits often. When the animal comes within the communication range of the router, the wireless transmission module will be activated to upload the compressed video data to a router and release the storage space in the CF card, as illustrated in Figure 1.

Figure 3(a) shows the camera and sensor unit attached to a collar. There are two sets of batteries, one on each side of the collar. Inside the collar, there is an inner tube which is filled with air and the air pressure is adjusted such that the collar can be comfortably fitted onto the animal's neck, as shown in Figure 3(b). At the top of the collar, there is a drop-off unit which is controlled by the drop-off control unit as shown in Figure 2. The drop-off control unit has a timer which allow us to set a specific time for the drop-off collar to release. The drop-off can be also triggered by other conditions, such as end of battery lifetime. To trigger the drop-off, the control unit generates a high-voltage signal which causes a squib inside the drop-off unit to fire. The gas generated when the squib fires moves a bolt inside the unit and releases the collar. Once dropped off, the collar sends out a radio signal repeatedly so that we can track it in the field and retrieve the system.

Another technical challenge in DeerCam system design is housing the system. A rugged housing is critical for protecting the system from damages, moisture, water, and extreme weather conditions. Wildlife species are often aggressive. They may smash or rub the system against trees. In our DeerCam system design, we have paid special attention to rugged housing design, as shown in Figure 3. Another important issue is whether the DeerCam system will affect the animal's health and change its behavior. According to recent studies [15], if the animal-mounted device weighs less than 3% of the animal body weight, the device won't have an negative impact on its health. Our research animals weigh approximately 160 pounds. Our current DeerCam system shown in Figure 3(a) weighs about four pounds, which is well below the weight threshold. Our research has not revealed changes in the behavior pattern of the animal.

III-B. Duty Cycle-Based Operation

Video encoding is energy-consuming. The limited battery power and the long lifetime requirement pro-

hibit us from continuous video monitoring. In addition, from the wildlife behavior analysis perspective, there is also no need to collect video continuously for 6 months. Therefore, we use a duty cycle-based operation scheme in the DeerCam system, as illustrated in Figure 4(a). Once the video sensor (Star-gate) is awake, it captures and encodes a short video segment (for example, 30 seconds) about the animal's activity, and goes back to the sleep state again. Each encoded video clip is called a *video sample*. The duty cycle can be characterized by three major parameters: average video sample duration δ , average sampling interval Δ , and picture quality D . In video compression, the picture quality is often measured by MSE (mean squared error) distortion.

We observe that configuration of these three duty-cycle parameters should be event-driven and should depend on the current activity state of the animal. In this project, we use a multi-level sensor-driven duty-cycle design, as illustrated in Figure 5(b). More specifically, the low-power Mote sensor in our DeerCam system operates continuously, collecting acceleration, audio, and GPS data. The multi-level duty cycle has four major levels. At the bottom level, there is a baseline duty cycle which wakes up the video sensor on a regular basis (for example, sun rise and sunset) or based on a preset schedule. On top of this, there are three levels of duty cycles triggered by acceleration, audio, and GPS sensors. From the temporal pattern of

Figure 3. (a) Configuration of the DeerCam system; (b) a DeerCam system deployed on a deer.

acceleration data, we can tell the major activity types, such as bedding/standing, walking, running, and eating, as shown in Figure 5(a). A sudden change in the audio signal, for example, animal screaming, often indicates an interesting event. From the GPS data, we can tell if the animal approaches interesting habitats or areas. From this set of sensor data, we try to estimate what the animal is doing and determine if this is an interest event. If so, the video sensor will be waked up to capture a video sample about the animal activity and its environment context. Based on the animal activity and event types, we can configuration of those three duty-cycle parameters (δ , Δ , D) accordingly. For example, when the animal is sleeping or resting, we can choose a very low sampling frequency; but when the animal is feeding or interacting with others, we would like to increase the sampling frequency, the sample duration, as well as the picture quality, if needed, to gather more detailed information. In Section IV-C-b, we will discuss how to configure these duty-cycle parameters so as to maximize the overall utility function under system resource constraints.

III-C. Energy Characterization of the DeerCam System

To optimize the energy consumption of DeerCam, the first step is to characterize and understand the energy consumption behavior of the system. Specifically, we need to know the energy consumption levels of major system components, such as the embedded CPU, mote sensor board, USB camera, and storage devices, at different system operation modes. To do this, we use the power measurement setup depicted in Figure 6(a) and follow a power consumption measurement procedure described in [28]. The operating voltage of the Star-gate 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 the remaining idle components on the Star-gate, 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. According to our experience, for this type of wildlife activity monitoring videos with a size of CIF (352×288) at 7–10 frames per second (fps), the

Figure 4. Duty cycle of video encoding and wireless transmission.

Figure 5. (a) An example of acceleration data collected by DeerCam; (b) Multilevel duty cycle design.

average bit rate is about 400 kbps². 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. Figure 6(b) shows the fractions of energy consumption for major system components. Here, we assume that a low-power USB camera from Logitech with a current draw of about 73 mA is used. It can be seen that the video encoder (processor) consumes a significant portion of the total energy and wireless transmission energy is about one quarter of the encoding energy due to delay-tolerant short-range wireless data transmission. Experimental studies in the literature (e.g. [29]) also show similar energy consumption behaviors of portable video devices in other communication settings, for example, live video streaming over wireless LAN.

IV. Power-Rate-Distortion Analysis for Video Encoding Energy Minimization

As discussed in Section III, most of the computational resources of the DeerCam system are used by the video encoder. Therefore, minimizing the video encoding energy plays a critical role in maximizing the operational lifetime of the DeerCam system. In this section, we present a power-rate-distortion (P-R-D) approach to video encoding energy minimization.

To reduce the energy consumption of video encoders, a number of algorithms, software and hardware energy-minimization techniques, including low-complexity encoder design [30–32], low-power embedded video encoding [33, 34], adaptive power control [29, 35, 36], and joint encoder and hardware adaptation [37, 39, 40] have been developed. These algorithms focus 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 saving and encoding performance [19]. In addition, even with existing energy saving technologies, the operational lifetime of portable video electronics is still very short, which

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

has become one of the biggest impediments to our technology future. Therefore, developing new energy optimization methods has become a urgent task.

Recently, we have been exploring a new approach to video encoding energy minimizing using P-R-D analysis and optimization. The P-R-D analysis tries to answer the following important question: *what is the minimum coding distortion (or maximum video quality) that we can achieve under bit rate and power constraints, and how we can achieve it?* We recognize that, from an information-theoretic perspective, incorporating power consumption into classical R-D analysis for generic source coding is a very difficult problem. Therefore, we choose to attack this problem from a design-analysis-optimization perspective. To successfully attack this problem, the following four major issues need to be carefully addressed: (1) The design problem—how to design an energy-scalable video encoder so that we can control its energy consumption? (2) The modeling problem—how to model its P-R-D behavior? (3) The optimization problem—what is the minimum power consumption for a video encoder to achieve a target R-D performance, or what is maximum R-D performance of a video encoder under a power consumption constraint? (4) The control problem—how to accurately control the video encoder to achieve the minimum energy consumption or maximum R-D performance? In the following sections, we explain these four problems in more detail and present our major approaches to addressing these problems.

IV-A. Energy-Scalable Video Encoder Design and Operational P-R-D Analysis

Our basic approach to designing an energy-scalable video encoder is to introduce a set of complexity control parameters $\{\gamma_1, \gamma_2, \dots, \gamma_L\}$ to control (scale down) the computational complexity of major encoding operations of the video encoder. With DVS

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

(Dynamic Voltage Scaling)³, a recently developed power control technology for microprocessors [21], this complexity-scalable video encoder can be translated into an energy-scalable video encoder [19].

More specifically, 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 $\Gamma = [\gamma_1, \gamma_2, \dots, \gamma_L]$ to control the power consumption of these modules. Therefore, the encoder complexity C is then a function of these control parameters, denoted by $C(\gamma_1, \gamma_2, \dots, \gamma_L)$. Within the DVS (dynamic voltage scaling) design framework [21], the microprocessor power consumption, denoted by P , is a function of computational complexity C , therefore, also a function of Γ , denoted by

$$P = \Phi(C) = P(\gamma_1, \gamma_2, \dots, \gamma_L), \quad (1)$$

where $\Phi(\cdot)$ is the power consumption model of the microprocessor [21]. For example, according to our measurement, the power consumption model of the Intel PXA255 XScale processor used in DeerCam is shown in Figure 7 (solid line). It can be well approximated by the following expression

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

where β is a constant. In the second step, we execute the video encoder with different configurations of complexity control parameters and obtain the corresponding R-D

Figure 7. Power consumption model with DVS.

³A number of sold processors, including the Intel XScale microprocessor used in many PDAs [38], support this DVS power control feature.

data, denoted by $D(R; \gamma_1, \gamma_2, \dots, \gamma_L)$. Note that this step is computationally intensive and is intended for offline analysis to obtain the P-R-D model only.

In the third step, we perform optimum configuration of the power 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:

$$\begin{aligned} \min_{\{\gamma_1, \gamma_2, \dots, \gamma_L\}} \quad & D = D(R; \gamma_1, \gamma_2, \dots, \gamma_L), \\ \text{s.t.} \quad & P(\gamma_1, \gamma_2, \dots, \gamma_L) \leq P, \end{aligned} \quad (3)$$

where P is the available power consumption for video encoding. Given the R-D data set $\{D(R; \gamma_1, \gamma_2, \dots, \gamma_L)\}$, 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 $\{\gamma_i^*(R, P)\}$, $1 \leq i \leq L$. Figure 8 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 [22]. 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.

From those experimental results obtained in operational P-R-D analysis, we can see that the P-R-D functions have an exponential behavior. For the convenience of analysis, we propose to approximate them using the following model

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

Here, σ^2 represents the variance of encoded picture. If it is a motion predicted video frame, σ^2 is the variance of the difference 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

$$g(P) = P^{\frac{1}{\gamma}}. \quad (5)$$

Here, $\gamma = 2.5$ for the Stargate microprocessor. For other microprocessors with DVS capabilities, we typically have $1 \leq \gamma \leq 3$ [21].

To use this P-R-D model for real-time video encoding, there are two major issues that need to be addressed.

First, we need to estimate the P-R-D model parameter λ in (4) from real-time video encoding statistics. Second, we need to determine the optimum complexity control parameters $\{\gamma_i^*(R, P)\}$, $1 \leq i \leq L$ of the video encoder for the input video segments. In our previous work [19], we observe that the P-R-D behavior of a video segment is closely related to its scene activity level, which is characterized by several feature variables, such as variance of motion vectors, Intra macroblock ratio, variance of difference pictures, etc. Based on this observation, we have developed a low-complexity scheme to estimate λ from motion statistics which can directly collected from the video encoder. To determine optimum complexity control parameters for a given input video sequence, we use a learning and classification approach. More specifically, we set a large set of training video sequences, which have a wide range of scene activity characteristics. Using the operational P-R-D analysis procedure, we obtain the P-R-D model, i.e., the minimum distortion function $D(R, P)$ and the optimum complexity control parameters $\{\gamma_i^*(R, P)\}$ for each video sequence. We then classify the training video sequences according to their scene activity levels into several (say, 5–7) clusters. Since the video sequences with each cluster have similar P-R-D behaviors, we can compute the average of their optimum complexity control parameters for each cluster and store it into a database. During real-time video encoding, we first classify the input video segment according to its scene activity level and determine its cluster. We then use the average optimum complexity control parameters of this cluster to control the input video segment.

IV-B. P-R-D Based Energy Minimization

Once the P-R-D model of the video encoder and the corresponding encoder complexity control parameters are

obtained, we can then use this model to minimize the energy consumption of video encoding. By incorporating the third dimension of power consumption into conventional R-D analysis, the P-R-D analysis gives us one extra dimension of flexibility in resource allocation and opens up a host of new opportunities for energy minimization. In the following, we use two examples, (a) trade-off between computation and communication and (b) trading bits for joules (energy), to demonstrate how the P-R-D model can be used for energy minimization of portable video devices.

IV-B-a. Trade-off between Computation and Communication

The P-R-D analysis enables us to explore the energy trade-off between computation (video encoding) and communication (wireless transmission) so as to achieve significant energy saving gain. According to the P-R-D model, 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 $P_t = c_t \cdot R$, 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 [44]. This is a simplified model to demonstrate the energy tradeoff between video encoding and wireless transmission. In our case, it is reasonable because 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 . This implies more power is needed for wireless data transmission. As illustrated in Figure 9, this leads to an energy tradeoff between video encoding and wireless data

Figure 8. Operational P-R-D functions of "Foreman" (a) and "Football" (b) CIF videos encoded with MPEG-4 at 10 fps.

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 the encoding power to an optimum level O , as illustrated in Figure 9 with a triangle, to minimize the overall power consumption. Our initial study shows that by exploring this type of energy trade-off between computation and communication, we are able to reduce the overall energy consumption by 30–60%, depending on the input video characteristics and wireless transmission condition. In practice, we need to consider fading channel and time-varying characteristics of input videos. For detailed treatment of this issue, see [41].

IV-B.b. Trading Bits for Joules

Our second observation is that, within the P-R-D analysis framework, the encoding power P is coupled with the encoding bit rate R . More specifically, for a given distortion D , the encoding power is a function of encoding bit rate, denoted by $P = P(R, D)$. This implies that we can minimize the total power consumption of video encoding by performing optimum bit allocation, just like in optimum bit allocation in conventional R-D analysis and optimization. To further illustrate this idea, let us consider the following example. Suppose we want encode a long video sequence on the DeerCam or a PDA with a bit budget R_T (say 2G bytes) at coding distortion D_0 (say 34 dB). We partition the video sequence into N segments $\{V_n | 1 \leq n \leq N\}$. Let R_n and P_n be the encoding bits and power consumption used by V_n . According to the P-R-D model, we have $P_n = P_n(D_0, R_n)$. Then, the energy minimization problem can be formulated as

$$\min_{\{R_n\}} P = \sum_{n=1}^N P_n(D_0, R_n), \quad \text{s.t.} \quad \sum_{n=1}^N R_n \leq R_T. \quad (6)$$

Our major result can be summarized as follows: if the input video data is non-stationary with time-varying scene characteristics, the P-R-D model enables us to explore the input source diversity, trade bits for joules (energy) between different video segments as outlined in (6), and 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 patterns. Therefore, the video sequence to be encoded by the portable device is often highly non-stationary. 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 [42], we demonstrate that, using this type of P-R-D based energy minimization, we are able to save the video encoding energy by up to 50%, depending on the non-stationary scene characteristics of the input videos. For a detailed treatment of this issue, see [42].

IV-C. Additional Energy

Minimization Schemes for DeerCam

Besides those two energy minimization schemes discussed in the previous section, we can also explore additional approaches for energy minimization based on the P-R-D analysis framework. In the following, we present two basic ideas to be further explored in our next step of research: (a) cross-layer energy minimization and (b) utility-driven resource allocation and energy minimization.

IV-C.a. Cross-Layer Energy Minimization

Cross-layer energy minimization is able to achieve significant energy saving by exploring joint design of hardware scheduling, encoder P-R-D control, and wireless transmission [37, 39, 40, 43]. The P-R-D model links the encoder resource parameters at the application layer with other resource parameters at the physical and link layers, which enables us to develop cross-layer energy minimization schemes. In cross-layer energy minimization for portable video communication systems, we need to consider the mobility of portable devices, the time-varying characteristics of wireless fading channels, the queuing behaviors of the processing (encoding) and transmission buffers, the energy allocation between video encoding and wireless transmission, and their joint impact to the end-to-end video quality, as well as to the overall energy saving performance.

Figure 9. Energy tradeoff between video encoding and wireless data transmission.

IV-C.b. Utility-Driven Resource Allocation and Energy Minimization

The second idea that can be explored for energy minimization in practical system design is utility-driven resource allocation and energy minimization. In practical applications, such as video surveillance, remote monitoring, and environmental tracking applications, the task of video encoding and streaming over portable devices is to collect important visual information about the targets, events, and their environmental contexts. Therefore, the performance of the portable video communication system should be measured by the amount of useful information that has been collected by the portable device, i.e., the utility of the video data. For example, the performance of the DeerCam system should be measured by the amount of visual information that is useful for wildlife behavior analysis. In the literature [6], network capacity and operation lifetime are two major performance metrics that are extensively used for energy minimization. However, within the context of video surveillance and monitoring, these two performance metrics are not sufficient. The amount of useful visual information, or the utility of video data, cannot be simply measured by its data size (in bits) or time duration (operational lifetime). A portable device may spend all of its energy on processing and transmitting a lot of bits which however contain little useful information for practical purposes.

Therefore, it is critical to perform utility-driven resource allocation and energy minimization such that the portable device spends its energy and other resources intelligently. To this end, we need to address the following two major issues. First, we need to establish a utility function which describes the inherent relationship between the resource control parameters and the overall utility. Second, based on this utility function, we can optimally configure the resource control parameters to maximize the utility function under the resource constraint. As a variation of this optimization procedure, we can also minimize the energy consumption while still satisfying the utility requirement. In the following, we use the DeerCam as an example to discuss how we could address these two issues.

As discussed in Section III-B, the DeerCam operates on a duty cycle, which is controlled by three major parameters, (δ, Δ, D) . We can see that these three parameters have a direct impact on resource consumption. For example, if we increase the sample duration δ , reduce the sampling interval Δ , or decrease the video distortion D (or equivalently increase the video quality), the video encoder will use more energy and generate more bits. Therefore, we also refer to them as resource control parameters. The overall utility of the collected video samples is a function of these three resource control parameters, denoted by $U(\delta, \Delta, D)$.

To set up a utility function $U(\delta, \Delta, D)$ for wildlife monitoring, we propose to use a procedure similar to perceptual video quality evaluation. More specifically, we capture a set of original video samples using animal-mounted video cameras, which cover a wide range of habitats and plant species. We then each original video sample with different configuration of resource control parameters (δ, Δ, D) . We organize a panel of wildlife experts and develop a video evaluation protocol to present each processed video sample to the expert panel and ask them to score from 0 to 10 the utility of each processed video sample for wildlife behavior analysis purposes. Based on scores provided by the expert panel, using statistical modeling, we can estimate the expression of $U(\delta, \Delta, D)$ and establish a utility function. It should be noted that this utility function is fully application-specific and knowledge-driven.

Meanwhile, we analyze the relationship between energy consumption and these resource control parameters. We call this as the energy consumption function, denoted by $E(\delta, \Delta, D)$. Now the utility-driven energy minimization problem becomes: optimally configuring the resource control parameters to minimize the energy consumption while satisfying the utility requirement, which can be mathematically formulated as:

$$\min_{(\delta, \Delta, D)} E(\delta, \Delta, D), \quad s.t. \quad U(\delta, \Delta, D) \geq U_0, \quad (7)$$

where U_0 is the minimum utility requirement. We expect that this utility-driven energy minimization approach will significantly reduce energy consumption.

V. Concluding Remarks

The capability of seeing what animal sees in the field is very important for wildlife monitoring and research. In this paper, we have introduced the DeerCam system, an integrated video and sensor system for animal-mounted wildlife monitoring. From the video and sensor data collected by DeerCam, wildlife researchers will be able to extract a wealth of scientific data for studying the behavior patterns of wildlife species and understanding the dynamic of wildlife systems. We have presented the system architecture of DeerCam, explained our system design goals and discussed major design issues. One of the central challenges in DeerCam system design is energy minimization. We have presented a new approach for energy minimization of portable video devices: P-R-D analysis and optimization. Results demonstrate that by incorporating the third dimension of power consumption into conventional R-D analysis, P-R-D analysis gives us one extra dimension of flexibility in resource allocation and energy minimization and allows us to significantly reduce energy consumption.

Acknowledgment

This work was supported in part by National Science Foundation under grant DBI-0529082.

References

- [1] L.D. Mech and S.M. Barber, *A critique of wildlife radio-tracking and its use in national parks: a report to the U.S. National Park Service*, U.S. Geological Survey, Northern Prairie Wildlife Research Center, Jamestown, N.D. 78 pages, 2002.
- [2] J.J. Millsaugh and J.M. Marzluff (editors). *Radio Tracking and Animal Populations*, Academic Press, San Diego, California, USA, 2001, 467 pages.
- [3] L.D. Mech, *A Handbook Of Animal Radio-tracking*, Univ. of Minn. Press, p. 108, 1983.
- [4] NASA Satellite Tracking of Threatened Species. http://sdcd.gsfc.nasa.gov/ISTO/satellite_tracking/, 2002.
- [5] P. Juang, H. Oki, Y. Wang, M. Martonosi, L.-S. Peh, and D. Rubenstein, "Energy-efficient computing for wildlife tracking: Design tradeoffs and early experiences with Zeburanet," *Tenth International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS-X)*, San Jose, CA, Oct. 5–9, 2002.
- [6] I.F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless sensor networks: a survey," *Computer Networks*, vol. 38, no. 4, pp. 393–422, Mar. 2002.
- [7] H. Gharavi and K. Ban, "Vision-based ad-hoc sensor networks for tactical operations," *World Wireless Congress, 3G Wireless 2002*, San Francisco, May 2002.
- [8] B. Thorstensen et al. "Electronic Shepherd: A Low-Cost, Low-Bandwidth, Wireless Network System," in *Proc. Second Intl. Conference on Mobile Systems, Applications and Services*, June 2004.
- [9] J.M. Kahn, R.H. Katz, and K.S.J. Pister, "Mobile Networking for Smart Dust," *ACM/IEEE Intl. Conf. on Mobile Computing and Networking (MobiCom 99)*, Seattle, WA, Aug. 17–19, 1999.
- [10] G. Pottic and W. Kaiser, "Wireless integrated network sensors," *Communications of the ACM*, vol. 43, no. 5, pp. 51–58, May 2000.
- [11] UC Davis Wildlife Health Center, Southern California Puma Project, <http://www.vetmed.ucdavis.edu/whc/scp/>, 2004.
- [12] R. Szewczyk, E. Osterwell, J. Polastre, M. Hamilton, A. Mainwaring, and D. Estrin, "Habitat monitoring with sensor networks," *Communications of ACM*, vol. 47, no. 6, p. 3440, 2004.
- [13] R. Szewczyk, A. Mainwaring, J. Polastre, and D. Culler, "An Analysis of a Large Scale Habitat Monitoring Application," in *Proceedings of the Second ACM Conference on Embedded Networked Sensor Systems (SenSys)*, Nov. 3–5, 2004.
- [14] K. Mayer, K. Taylor, and K. Ellis, "Cattle Health Monitoring Using Wireless Sensor Networks," *The 2nd IASTED International Conference on Communication and Computer Networks*, Cambridge Massachusetts, Nov. 8–10, 2004.
- [15] J. Beringer, J.J. Millsaugh, J. Sartwell, and R. Woeck, "Real-time video recording of food selection by captive white-tailed deer," *Wildlife Society Bulletin*, vol. 32, no. 3, 2004.
- [16] T.L. Cutler and D.E. Swann, "Using remote photography in wildlife ecology: a review," *Wildlife Society Bulletin*, no. 27, pp. 571–581, 1999.
- [17] D.I. King, R.M. DeGraaf, P.J. Champlin, and T.B. Champlin, "A new method for wireless video monitoring of bird nests," *Wildlife Society Bulletin*, no. 29, pp. 349–353, 2001.
- [18] A.B. Cooper and J.J. Millsaugh, Accounting for variation in resource availability and animal behavior in resource selection studies. Pages 243–274 in J. J. Millsaugh and J. M. Marzluff, editors. *Radio Tracking and Animal Populations*. Academic Press, San Diego, California, USA, 2001.
- [19] Z. He, Y. Liang, L. Chen, I. Ahmad, and D. Wu, "Power-rate-distortion analysis for wireless video communication under energy constraint," *IEEE Transactions on Circuits and Systems for Video Technology*, May, 2005.
- [20] T. Wiegand, "Text of Committee Draft of Joint Video Specification (ITU-T Rec. H.264—ISO/IEC 14496-10 AVC)," *Document JVTG167, 3rd JVT Meeting*, Fairfax, Virginia, USA, May 6–10, 2002.
- [21] T. Burd and R. Broderson, "Processor Design for Portable Systems," *Journal of VLSI Signal Processing*, vol. 13, no. 2, pp. 203–222, Aug. 1996.
- [22] B. Zeng, R. Li, and M.L. Liou, "Optimization of fast block motion estimation algorithms," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 7, no. 6, pp. 833–844, Dec. 1997.
- [23] H.D.J.N. Aldridge and R.M. Bringham, "Load carrying and maneuverability in an insectivorous bat: A test of the 5% rule of radio-telemetry," *Journal of Mammalogy*, vol. 69, pp. 379–382.
- [24] J.C. Withey, et al, "Effects of tagging and location error in wildlife radiotelemetry studies," *Radio Tracking and Animal Populations* (Millsaugh, J.J. and Marzluff, J.M., eds), pp. 43–75, Academic Press, 2001.
- [25] J.R. Fischer, L.H. Creekmore, R.L. Marchinton, S.J. Riley, S.M. Schmitt, and E.S. Williams, "External review of chronic wasting disease management in Wisconsin," Oct. 10, 2003.
- [26] D. Estrin, R. Govindan, J. Heidemann, and S. Kumar, "Next century challenges: Mobile networking for smart dust," in *ACM MOBICOM*, Seattle, WA, Aug. 1999.
- [27] http://www.xbow.com/Products/Product_pdf_files/Wireless_pdf/Stargate_Datasheet.pdf
- [28] C.B. Margi, V. Petkov, K. Obraczka, and R. Manduchi, "Characterizing energy consumption in a visual sensor network testbed," *Proceedings of 2nd International Conference on Testbeds and Research Infrastructures for the Development of Networks and Communities (TRIDENTCOM)*, Mar. 2006, pp. 8–15.
- [29] X. Lu, Y. Wang, and E. Erkip, "Power efficient H.263 video transmission over wireless channels," *Proceedings of 2002 International Conference on Image Processing*, Rochester, New York, Sept. 2002.
- [30] K. Hyungjoon, N. Kamaci, and Y. Altunbasak, "Low-complexity rate-distortion optimal macroblock mode selection and motion estimation for MPEG-like video coders," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 15, no. 7, pp. 823–834, July 2005.
- [31] I.M. Pao and M.T. Sun, "Statistical Computation of Discrete Cosine Transform in Video Encoders," *Journal of Visual Communication and Image Representation*, vol. 9, no. 2, pp. 163–170, June 1998.
- [32] D.S. Turaga, M. van der Schaar, and B. Pesquet-Popescu, "Complexity scalable motion compensated wavelet video encoding," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 15, no. 8, pp. 982–993, Aug. 2005.
- [33] J. Chen and K.J.R. Liu, "Low-power architectures for compressed domain video coding co-processor," *IEEE Transactions on Multimedia*, vol. 2, no. 2, pp. 111–128, June 2000.
- [34] J. Villaseñor, C. Jones, and B. Schoner, "Video Communications using Rapidly Reconfigurable Hardware," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 5, pp. 565–567, Dec. 1995.
- [35] P. Agrawal, J.-C. Chen, S. Kishore, P. Ramanathan, and K. Sivalingam, "Battery power sensitive video processing in wireless networks," *Proceedings IEEE PIMRC'98*, Boston, Sept. 1998.
- [36] W.P. Burleson, P. Jain, and S. Venkatraman, "Dynamically Parameterized Architecture for Power-Aware Video Coding: Motion Estimation and DCT," *Proceedings of the Second USF International Workshop on Digital and Computational Video*, 2001.
- [37] V. Akella, M. van der Schaar, and W.-F. Kao, "Proactive Energy Optimization Algorithms for Wavelet-Based Video Codecs on Power-Aware Processors," *IEEE International Conference on Multimedia and Expo*, pp. 566–569, July 6–8, 2005.
- [38] Intel Inc, "Intel XScale Technology," <http://www.intel.com/design/intelxscale>.
- [39] D.G. Sachs, W. Yuan, C.J. Hughes, A.F. Harris, S.V. Adve, D.L. Jones, R.H. Kravets, and K. Nahrstedt, "GRACE: A Cross-Layer Adaptation Framework for Saving Energy," Sidebar in *IEEE Computer, special issue on Power-Aware Computing*, Dec. 2003, pp. 50–51.
- [40] D.G. Sachs, S. Adve, and D.L. Jones, "Cross-layer Adaptive Video Coding to Reduce Energy on General-Purpose Processors," *Proceedings of the International Conference on Image Processing, 2003 (ICIP '03)*, Barcelona, Spain, Sept. 2003.
- [41] Z. He and D. Wu, "Resource allocation and performance limit analysis of wireless video sensors," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 16, no. 5, pp. 590–599, May 2006.
- [42] W. Cheng, X. Chen, and Z. He, "Doubling of the operational lifetime of portable video communication devices using power-rate-distortion analysis and control," *Proceedings of International Conference on Image Processing*, Atlanta, GA, Oct. 2006.
- [43] V. Akella, M. van der Schaar, and W.-F. Kao, "Proactive Energy Optimization Algorithms for Wavelet-Based Video Codecs on Power-Aware Processors," *IEEE International Conference on Multimedia and Expo*, July 6–8, 2005, pp. 566–569.
- [44] T.S. Rappaport, *Wireless Communications: Principles and Practice*, Prentice Hall, New Jersey, 1996.

Zhihai He 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. 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. His current research interests include image/video processing and compression, network transmission, wireless communication, computer vision analysis, sensor network, and embedded system design. 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.

Jay Eggert is a Ph.D. candidate in electrical and computer engineering at the University of Missouri. He received his B.S. degree from Purdue University in electrical engineering and was previously employed by Science Applications International Corporation (SAIC) as a Senior Engineer. His current research interests include application of image/video processing and compression to remote systems, wireless communication, sensor networks, and embedded system design.

Wenye Cheng received his B.S. degree in infrared technologies from Xidian University, Xi'an, China in 1983 and M.S. degree from Electrical and Computer Engineering of University of Missouri, Columbia in 2006. He is currently with Harmonic Inc. His research interests include embedded system design for image/video compression, processing, and communication.

Xiwen Zhao received the B.E. degree in thermal engineering from Tsinghua University, Beijing, China, in 1998, the M.E. degree in electrical engineering from the Graduate University of Chinese Academy of Sciences, Beijing, China, in 2002, the M.S. degree in electrical and com-

puter engineering from the University of Minnesota, Minneapolis, in 2006. He is now pursuing the Ph.D. degree in electrical and computer engineering at the University of Missouri, Columbia. He is interested in image/video processing and compression.

Joshua Millspaugh received his B.S. degree in Forest Biology from the State University of New York in Syracuse, New York, a M.S. degree from South Dakota State University in Wildlife Ecology, and a Ph.D. from the University of Washington in Wildlife Ecology in 1999. Joshua joined the Department of Fisheries and Wildlife Sciences as an Assistant Professor in 1999. He is currently an Associate Professor. He has received the Wildlife Society Award for Best Article, the University of Missouri Provost Outstanding Junior Faculty Teaching Award, and the University of Missouri William T. Kemper Fellowship for Excellence in Teaching. His current research interests includes the design and analysis of radio tracking studies, application of sensor technology in ecology, vertebrate population ecology, and the refinement and use of hormone assays in wildlife conservation.

Remington Moll received his B.S. summa cum laude from the University of Missouri in Fisheries and Wildlife Science in 2006, where he is currently an M.S. candidate in Fisheries and Wildlife Science. His research focuses on animal-borne video systems, wildlife stress physiology, quantitative ecology and animal behavior.

Jeff Beringer is a resource scientist with the Missouri Department of Conservation (MDC). He received a B.S. in wildlife management from the University of Wisconsin-Stevens Point and a M.S. in wildlife and fisheries science from the University of Tennessee. His primary interest is deer and furbearer research and management.

Joel Sartwell is a systems analyst with the Missouri Dept. of Conservation. He received a BS in chemistry from the University of Missouri-Kansas City and a Ph.D. in physical chemistry from the University of Nebraska. His interests include scientific software development, user interfaces, simulation, and statistical wildlife biology. His current projects include data capture systems for hand-held computers, data manipulation and standardization with xml, and radio-tracking analysis.