Resource Allocation and Performance Analysis of Wireless Video Sensors

Zhihai He, *Senior Member, IEEE*, and Dapeng Wu, *Senior Member, IEEE*

***Abstract*—Wireless video sensor networks (WVSNs) have been envisioned for a wide range of important applications, including battlefield intelligence, security monitoring, emergency response, and environmental tracking. Compared to traditional communication system, the WVSN operates under a set of unique resource constraints, including limitations with respect to energy supply, on-board computational capability, and transmission bandwidth. The objective of this paper is to study the resource utilization behavior of a wireless video sensor and analyze its performance under the resource constraints. More specifically, we develop an analytic power-rate-distortion (P-R-D) model to characterize the inherent relationship between the power consumption of a video encoder and its rate-distortion performance. Based on the P-R-D analysis and a simplified model for wireless transmission power, we study the optimum power allocation between video encoding and wireless transmission and introduce a measure called achievable minimum distortion to quantify the distortion under a total power constraint. We consider two scenarios in wireless video sensing, small-delay wireless video monitoring and large-delay wireless video surveillance, and analyze the performance limit of the wireless video sensor in each scenario. The analysis and results obtained in this paper provide an important guideline for practical wireless video sensor design.**

***Index Terms*—Power consumption, rate-distortion, resource allocation, wireless video compression, wireless sensor network.**

I. INTRODUCTION

A WIRELESS sensor network is a system of geographically distributed sensor nodes that communicate with each other over a wireless medium. Without the need for a communication infrastructure, as is the case in the traditional cellular networks, the wireless sensor network is self-organized and highly dynamic, with each communication node serving as both server and router for data transmission [3], [26]. In a wireless video sensor network (WVSN), each sensor, which is equipped with video capture and processing capabilities, is tasked to capture digital visual information about target events or situations and deliver the video data to a remote control unit (RCU) for further information analysis and decision making, as illustrated in Fig. 1. Because of its unique features of rapid deployment, flexibility, low maintenance cost, and robustness, the WVSN has been envisioned for a wide range of important applications, including security monitoring, emergence response, environmental tracking, and health monitoring.

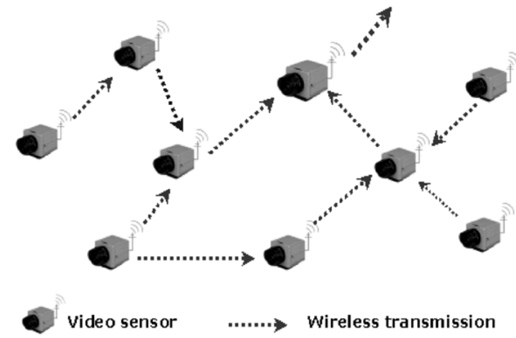

Fig. 1. Illustration of a WVSN.

The ultimate goal in communication system design is to optimize the system performance under resource constraints. In traditional video communication applications, such as digital TV broadcast, personal video recording, consumer video electronics, and video-on-demand, the major resource constraint is in the form of transmission bandwidth or storage space, which determines the output bit rate of a video encoder [13], [15], [16]. Computational complexity and energy supply are, in general, not a major concern, because the video encoding can be done offline on high-performance computers with a wired power supply. To analyze the signal processing and communication behavior of a video system under the bandwidth constraint, rate-distortion (R-D) theories have been developed, from early Shannon source coding theorem [6], [12] to recent R-D analysis of modern video coding systems [13], [15], [16], [30]. Based on the R-D models, rate allocation and control algorithms have been developed to control the encoding bit rate and optimize the system performance under bit-rate constraints [16], [30].

Compared to traditional communication systems, the WVSN operates under a set of unique resource constraints, including limitations with respect to energy supply, on-board computational capability, and transmission bandwidth. Therefore, there is a need to extend the traditional R-D analysis by considering these new resource constraints in the WVSN. The objective of this study is to analyze the impact of resource constraints on the performance limit of a wireless video sensor. The analysis and results developed in this study will answer the following fundamental question: *if a wireless video sensor is deployed with some resources (e.g., energy), what is the maximum video sensing performance that the sensor node can achieve?* In WVSNs, video compression and wireless transmission are the major operations on each wireless video sensor. For a battery-powered wireless

Manuscript received August 13, 2004; revised March 1, 2005, and July 12, 2005. This work was supported in part by the National Science Foundation under Grant DBI-0529082 and Grant DBI-0529012. This paper was recommended by Associate Editor H. Gharavi.

Z. He is with the Department of Electrical and Computer Engineering, University of Missouri, Columbia, MO 65211 USA (e-mail: HeZhi@missouri.edu).
D. Wu is with the Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL 32611 USA (e-mail: wu@ece.ufl.edu).

Digital Object Identifier 10.1109/TCSVT.2006.873154

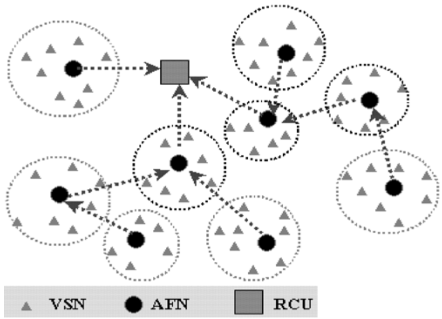

Fig. 2. Network architecture of the WVSN. VSN: video sensor node; AFN: aggregation and forwarding node; RCU: remote control unit.

video sensor, it is essential to maximize the power efficiency of these two operations. In this study, we will develop an analytic power-rate-distortion (P-R-D) model to characterize the inherent relationship between the power consumption of the video encoder and its R-D performance. We will analyze the power consumption in wireless transmission and its impact on the overall video quality. Based on these analytic resource utilization models, we will finally study the performance of wireless video sensors under resource constraints.

The remainder of this paper is organized as follows. In Section II, we present a reference architecture for WVSNs and discuss the major resource constraints. Section III presents the conceptual framework for resource allocation and performance optimization. The P-R-D analysis for a video encoder is presented in Section IV. Based on the P-R-D model and a simplified model for wireless video transmission, we study the performance limit of wireless video sensors and introduce the concept of achievable minimum distortion in Section V. In Section VI, we consider two scenarios of wireless video sensing and study the performance limit of wireless video sensors in a practical setting. Concluding remarks and future research directions are given in Section VII.

II. Wireless Video Sensor Networks (WVSN)

A typical wireless sensor network has three types of communication nodes: sensors, aggregation and forwarding nodes (AFNs), and an RCU. The AFNs aggregate the data collected by sensor nodes and forward it to the RCU, which is the destination of all video data. These communication nodes are typically organized into a tiered network architecture [21], [41], with AFNs at the top tier, each managing a cluster of video sensor nodes (VSNs) at the bottom tier, as illustrated in Fig. 2.

Existing research in wireless sensor networks has focused on the architecture design and performance optimization of conventional sensors, such as chemical, biological, and temperature sensors, in which data rates are often low, data processing is simple, and the power consumed in data processing is often negligible [3], [21], [35]. In contrast to this, WVSNs have their unique properties. More specifically, the data rate of raw videos generated by a video sensor is extremely high. This requires the video to be efficiently compressed before transmission; otherwise, the required network bandwidth and power consumption for wireless transmission are tremendous. However, efficient video compression often involves sophisticated and computationally intensive encoding operations. Therefore, the video encoder consumes a significant portion of the total energy supply on a wireless video sensor. Experimental studies show that, for relatively small picture sizes, such as QCIF (176×144) videos, video encoding (with H.263) consumes about two thirds of the total power in video communication over a wireless local area network (WLAN) [1], [28]. For pictures of higher resolutions, it is expected that the encoder power consumption will be more significant. This requires us to develop a framework to analyze the power consumption in both video encoding and wireless transmission and study the power allocation between these two operations.

III. Resource Allocation and Performance Analysis

The existing research on resource allocation and performance optimization for wireless sensor networks has focused on the data transmission over wireless networks. A variety of packet routing, flow and topology control, and power management schemes have been proposed to minimize the data transmission energy and optimize the system performance, such as network capacity or operational lifetime [3], [21], [35]. However, little research has been done to analyze the resource utilization behavior of individual sensor nodes and explore their performance limit. The major reason for this is that, in the conventional wireless sensor network design, the function of each sensor node is very simple, as it just transmits data or relays packets for others according to a routing protocol. The data processing on the sensor node is assumed to be very simple, basically, reading the sensor measurement, and the energy consumption in data processing is very little and is often assumed to be negligible. However, within the context of WVSNs, with the computationally intensive and energy-demanding video compression incorporated into the sensor node, its energy consumption behavior becomes very complex. Therefore, there is a need to analyze the energy consumption behavior of individual sensor nodes and optimize their performance under energy constraints. Since the sensor node is the basic operation unit of the WVSN system, its performance analysis is the first step, as well as the gateway, to the performance analysis and optimization of the whole sensor network. The analysis will also provide a solid foundation for protocol design and algorithm development at upper layers of the network system.

In this section, we outline the basic framework for resource allocation and performance analysis of the wireless video sensor. Note that, in the proposed system architecture for WVSNs, the major task of the wireless video sensor is to compress the video sensor data and send the compressed bit stream to the AFN. Its performance is measured by the quality of videos delivered to the AFN. In the literature, a commonly used measure for video quality is the end-to-end distortion, denoted by D, which is the mean square error (MSE)[1] between the original picture F captured by the sensor node and the received picture $\tilde{F}$ at the AFN. This distortion is denoted by $\mathrm{MSE}(F, \tilde{F})$. In video communication applications over wireless networks,

[1]Because of its mathematical tractability, we use MSE as the performance measure for video sensors. More sophisticated measures, such as those considering human perception [24], could be considered in future work.

the end-to-end distortion D consists of two parts: source coding distortion caused by lossy video compression, denoted by D_s, and transmission distortion caused by transmission errors, denoted by D_t. The specific definitions of D_s and D_t are given by

$$D_s = \mathrm{MSE}(F, \hat{F}) \qquad \text{and} \qquad D_t = \mathrm{MSE}(\hat{F}, \tilde{F}) \quad (1)$$

where $\hat{F}$ is the encoder reconstruction of the original picture F. Here, we assume that a hybrid video encoder, such as H.263 or MPEG-4, is used for video compression. It has been observed the encoding error and transmission error are uncorrelated, therefore, $D = D_s + D_t$ [17], [37]. Let P_0 be the power supply of the sensor node. Let P_s and P_t be the amount of power used for video encoding and data transmission. Assuming that video encoding and data transmission dominate in the power budget P_0, the following constraint is obvious:

$$P_0 = P_s + P_t. \quad (2)$$

Intuitively, the encoding distortion D_s is a function of P_s and source coding bit rate R_s, which is denoted by $D_s(R_s, P_s)$. The transmission distortion D_t is a function of the transmission power P_t and transmission distance d, which is denoted by $D_t(P_t, d)$. Therefore, conceptually, the resource allocation and performance optimization of the wireless video sensor can be formulated as follows:

$$\begin{aligned} &\min_{\{P_s, P_t, R_s\}} D = D_s(R_s, P_s) + D_t(P_t, d) \\ &\text{s.t. } P_s + P_t \le P_0. \end{aligned} \quad (3)$$

To obtain an optimized allocation of P_s, P_t, and R_s, we need to analyze the P-R-D behavior of the video encoder and obtain some function $D_s(R_s, P_s)$. This will be described in Section IV. In addition, we need to analyze the behavior of the transmission errors in video streaming over wireless networks and obtain some function $D_t(P_t, d)$. This will be achieved by two major steps. In the first step, we consider the ideal case where packet transmissions are free of errors and study the achievable minimum distortion of the wireless video sensor, which serves as a benchmark for the nonideal (error-prone) case. We will discuss the achievable minimum distortion in Section V. In the second step, we consider nonideal cases where transmissions are prone to errors. We study two different approaches to handle transmission errors and study the power allocation problem in (3). This will be discussed in Section VI.

IV. P-R-D Analysis

The main objective of this section is to analyze the energy consumption of video encoding and its impact on the R-D performance.

A. Related Work

Complexity-scalable algorithms have been developed in the literature for low-power video encoding [7], [31], [38]. Hardware implementation technologies have also been developed to improve the video coding speed [20], [38]. The power consumption behavior of video encoding has been experimentally evaluated in [1] and [28]. However, little research has been done to establish an analytical P-R-D model for complexity analysis, optimum power allocation, and control. Incorporating the power consumption into the existing R-D analysis framework is a challenging task, since the power consumption and R-D performance are two totally different concepts, and it is difficult to establish a direct relationship between these two in a typical video compression system that often has many sophisticated prediction and encoding operations [36], [39].

B. Parametric Power-Scalable Video Encoding Design and Analysis

To analyze and control the power consumption of a video encoder, we take two major steps. In the first step, we design a power-scalable video encoding architecture. In the second step, we analyze the relationship between the power consumption and R-D performance of the power-scalable video encoder. To dynamically control the power consumption of a microprocessor on a mobile device, a CMOS circuits design technology, called dynamic voltage scaling (DVS), has been recently developed [27], [29]. The central idea in DVS is that the power consumption of the microprocessor, as well as the processing speed, can be dynamically controlled by adjusting the voltage supply of the CMOS circuit. Various chip makers, including AMD [4] and Intel [23], have recently announced and sold processors with this power-scaling feature. From the video encoding perspective, with the DVS hardware technology, we can control the power consumption of a video encoder by adjusting its computational complexity. This effectively translates the complexity of a video encoding algorithm, denoted by C_s, into the power consumption of the video encoder, denoted by P_s. The relationship between C_s and P_s is given by the power consumption model of the microprocessor $P_s = \Phi(C_s)$ [23], [27]. We can see that, within the DVS framework, the power scalability is equivalent to complexity scalability.

In our previous work [18], [19], we successfully developed a parametric video encoding architecture that is fully scalable in computational complexity and power consumption. This is achieved by introducing a parameter set $\Gamma = [x, y, z]$ to control the computational complexity and the power consumption of the encoder. Mathematically, both P_s and C_s are functions of Γ, denoted by $P_s(\Gamma)$ and $C_s(\Gamma)$, respectively. By analyzing the R-D behavior of each control parameter, we have successfully obtained an analytic expression for the video encoding distortion D_s as

$$\begin{aligned} D_s = &D_s(R_s, \Gamma) = (1-z)^2(\beta_0 + \beta_1) \\ &+ 2(2z - z^2)(\beta_0 + \beta_1 e^{-\beta_2 x}) \\ &\cdot [\frac{1}{2}(1-y)^2 + y(1 + a_0 y) \cdot 2^{-2\gamma \frac{R_s}{y}}] \end{aligned} \quad (4)$$

where $\{\beta_0, \beta_0, a_0\}$ are model parameters estimated from previous coding statistics [19]. Let $\bar{P}_s$ be the amount of power supply allocated for video compression. To maximize its R-D performance under the power constraint, the encoder needs to

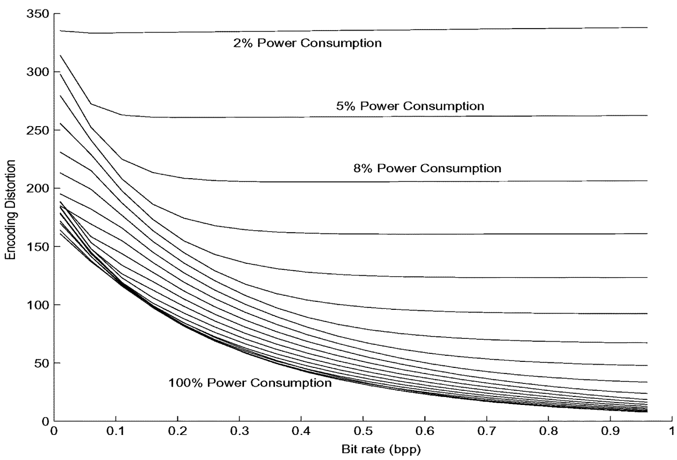

Fig. 3. R-D curves for different power consumption levels.

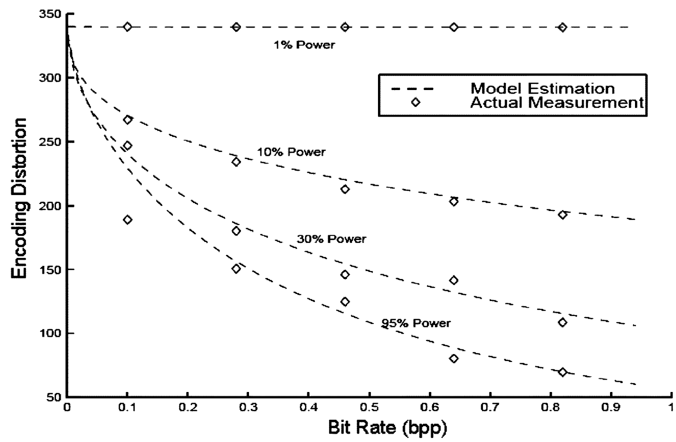

Fig. 4. R-D curves for different power consumption levels given by the analytic model (6).

perform an optimum allocation of the computing power (or processing cycles) among different encoding modules and find the best configuration of the complexity control parameters Γ, i.e.,

$$\min_{\Gamma} D_s(R_s, \Gamma), \quad \text{s.t.} \quad P_s(\Gamma) \le \bar{P}_s. \tag{5}$$

By solving this minimization problem with numerical methods, we have obtained the P-R-D model $D_s(R_s, P_s)$ [18], [19]. Fig. 3 shows the $D_s - R_s$ curves[2] at different power consumption levels $\bar{P}_s$ in percentages of $P_s^{\max}$. Here, $P_s^{\max}$ is the maximum power consumption level of the video encoder.

C. Analytic P-R-D Model

It should be noted that, in [18], we use a numerical procedure to find the minimum solution of (5). This is because the objective function in (4) is quite complicated, and it is impossible to obtain an analytical expression for the P-R-D function. However, an analytic P-R-D model is needed for resource allocation and performance analysis. In this study, we develop an analytic expression for the P-R-D model $D_s(R_s, P_s)$ based on the P-R-D behavior data we obtained in our previous work [18].

Several observations can be made from the P-R-D plots in Fig. 3.

1) When $P_s = 0$ and $R_s = 0$, the coding distortion should be equal to the variance of the input video, because, in this case, the encoder has no resource to perform any video compression and transmission operations.
2) As is suggested by the classical R-D models [6], the relationship between the coding bit rate R_s and distortion D_s is exponential. From Fig. 3, we can see that when the power supply level decreases, the $D_s(R_s)$ function becomes flatter, which means that the video compression efficiency is reduced. This is because the encoder has less power and computational capability to squeeze out the temporal and spatial redundancy in the input video data using motion prediction, spatial transform and data representation.

[2]It should be noted that here R_s includes both motion vector bits and transform coefficients coding bits. However, in [18], only the coefficients bits are counted in R_s.

Based on these two observations, we propose the following P-R-D model for power-scalable video encoding:

$$D_s(R_s, P_s) = \sigma^2 e^{-\gamma R_s \cdot g(P_s)}, \quad P_s \in [0,1], \qquad R_s \ge 0 \tag{6}$$

where σ^2 is the input variance, γ is a model parameter related to encoding efficiency. The function $g(\cdot)$ is the inverse function of the power consumption model $\Phi(\cdot)$ of the microprocessor. Comparing the P-R-D model in (6) against the traditional R-D model, we can see that $\gamma \cdot g(P_s)$ represents the coding efficiency of video encoder.

To test the accuracy of the P-R-D model in (6), we conduct the following experiment. The power consumption model used here is $P_s = \Phi(C_s) = C_s^{3/2}$ or $g(P_s) = P_s^{2/3}$ [23]. We run the complexity-scalable video encoder on a 400-MHz PDA with an Intel XScale microprocessor. The test video is "Foreman" QCIF (176×144) encoded at 10 frames per second (fps). We run the video encoder on various combinations of bit rates R_s and complexity control parameter sets Γ. For each set of complexity control parameters, we record the corresponding running time (in seconds), which represents the encoder complexity C_s. C_s is then translated into encoder power consumption P_s using the power consumption model $\Phi(\cdot)$. For a given R_s and P_s, we search for the minimum distortion D_s. In Fig. 4, we plot these points $[R_s, P_s, D_s]$ in diamonds. The dashed lines show the approximation with the analytical model in (6). It can be seen that the exponential model is fairly accurate. Simulations over other test video sequences yield similar results.

It should be noted that the P-R-D model in (6) is intended to characterize the P-R-D behavior of the encoder at relatively large scales, i.e., a group of video pictures (GOP), instead of a single video frame. In real-world applications, the energy supply of a wireless video sensor is expected to last for a relatively long period of time, e.g., hours or even days, and the average P-R-D modeling at the GOP level will be sufficiently accurate to capture the power consumption behavior of the video encoder for efficient power management. In practice, the model parameters γ can be obtained from the previous measurements of the coding distortion D_s, bit rate R_s, and power consumption P_s.

V. Achievable Minimum Distortion

In this section, based on the P-R-D model and a simplified power consumption model for wireless video transmission, we are going to study the power allocation on a wireless video sensor and explore its performance limit under the power constraint.

A. Achievable Minimum Distortion

As mentioned before, the energy supply of a sensor node is mainly used by video compression and wireless transmission. The following two observations can be made.

- **Case A:** If we decrease the encoder power consumption P_s, the distortion D_s increases. This is because the video encoder does not have enough computational resource and processing power to encode the video data, i.e., $P_s \downarrow \Rightarrow D_s \uparrow$.
- **Case B:** Since the total power consumption P_0 is fixed, and $P_0 = P_s + P_t$, if we increase P_s, then P_t decreases. This implies that less bits can be transmitted because the transmission energy is proportional the number of bits to transmit. Therefore, $P_s \uparrow \Rightarrow P_t \downarrow \Rightarrow R_s \downarrow \Rightarrow D_s \uparrow$.

It can be seen that, when the encoding power P_s goes too low or too high, the encoding distortion D_s will become large. This implies that there exists an optimal power P_s that minimizes the video distortion D_s. In the following, based on a *simplified* power consumption model for wireless transmission, we study the performance of wireless video sensors. More specifically, we assume that the transmission power is properly chosen such that the bit error rate (BER) at the receiver side is very low and the transmission errors can be neglected. Note that the transmission power is given by

$$P_t = \eta(d) \cdot R_s \quad \text{and} \quad \eta(d) = \omega + \gamma d^n \tag{7}$$

where R_s is the transmission bit rate, d is the transmission distance, and n is the path-loss exponent [5], [33]. Therefore

$$P_0 = P_s + P_t = P_s + \eta(d) R_s \tag{8}$$

and

$$R_s = \frac{P_0 - P_s}{\eta(d)}. \tag{9}$$

Since we assume that the transmission errors are negligible, we have $D_t = 0$ and $D = D_s$ According to the P-R-D model, we have

$$\begin{aligned} D &= D_s(R_s, P_s)|_{R_s = \frac{P_0 - P_s}{\eta(d)}} \\ &= D_s\left(\frac{P_0 - P_s}{\eta(d)}, P_s\right), \quad 0 \le P_s \le P_0. \end{aligned} \tag{10}$$

It can be seen that D is a function of P_s, denoted by $D(P_s)$. Using the analytic P-R-D model in (6), we compute the function $D(P_s)$ in (10) and plot it in Fig. 5. Here, the power supply of the wireless video sensor is $P_0 = 0.3$ W. This is a typical plot of $D(P_s)$. It can be seen that $D(P_s)$ has a minimum point,

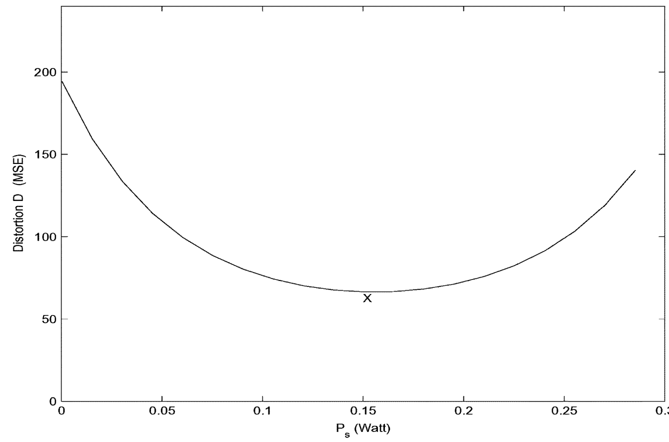

Fig. 5. Plot of $D(P_s)$ in (10) and illustration of the AMD, given fixed total power consumption P_0.

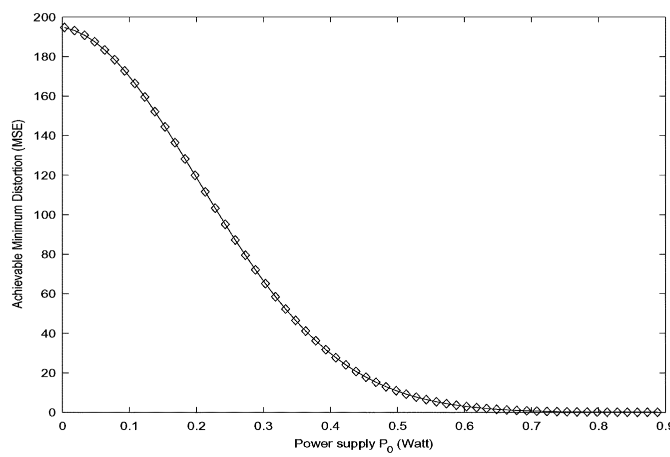

Fig. 6. Plot of AMD(P_0).

which is the minimum encoding distortion (or maximum video quality) that a wireless video sensor can achieve for a given power supply, no matter how the sensor node allocates its power resource between video encoding and wireless transmission. We call this minimum distortion as achievable minimum distortion (AMD). In Fig. 6, we plot the AMD as a function of the power supply P_0.

VI. Performance Analysis of the Wireless Video Sensor

It should be noted that the performance bound given by the AMD is not tight because the transmission error is not considered. In this section, we consider two different approaches to handle the transmission error and analyze the performance of wireless video sensors in each case.

Depending on the delay requirement, the WVSN applications can be classified into two basic categories. In the first category, which is WVSN monitoring, the video sensor data is required to be transmitted over the WVSN to the destination with a small delay for fast response and decision making. In the second category, which is WVSN surveillance, there is no stringent delay requirement. It only requires that the video sensor data be successfully delivered to the destination before the surveillance mission is completed. In this section, we consider two scenarios of wireless video sensing and investigate their AMD performance limits. In the first scenario of WVSN surveillance, where the transmission delay is not a major concern but high video quality is a top priority, the packet retransmission is used to assure that video packets are received correctly. In the second scenario of WVSN monitoring, which has a stringent delay requirement, the packet retransmission is not allowed. In this case, we are trying to use the transmission distortion model to control and minimize the end-to-end distortion. The performance analysis for these two scenarios are discussed in Section VI-A and B, respectively.

A. Performance Analysis for Large-Delay WVSN Surveillance

In video surveillance, e.g., aerial video surveillance, the video quality is of high priority. Therefore, it is often required that the video data be received correctly. In this scenario, we assume that packet retransmissions will be requested if the packet is received in error. Since the transmission energy is proportional to the number of bits that are transmitted, more packet retransmissions require more transmission energy. However, packet retransmissions increase the probability that the packet is received correctly. In the following, we analyze this tradeoff between power consumption and reliability and its impact on the performance of a wireless video sensor. More specifically, we first analyze the power consumption in wireless video transmission with packet retransmissions and determine the optimum transmission power. We then find that there is a unique behavior of the optimum transmission power, which enables us to apply the AMD analysis developed in Section V directly to performance analysis in the WVSN surveillance scenario.

Let E_t^b be the transmission energy per bit at the VSN, P_r be the received power at the AFN, and E_r^b be the received energy per bit at the AFN. According to [33], we have

$$E_r^b = h \times d^{-n} \times E_t^b \tag{11}$$

where d is the distance between the VSN and the AFN, n is the path-loss exponent (which is typically $n \in [2,5]$), and h is a parameter depending on the transmitter, the wireless channel, and the receiver. For example, in the case of free space transmission, we have $n=2$ and $h=(G_tG_r\lambda^2)/((4\pi)^2\kappa)$ [33], where G_t is the transmitter antenna gain, G_r is the receiver antenna gain, κ is the system loss factor not related to propagation $(\kappa \geq 1)$, and λ is the wavelength in meters.

If BPSK is used for modulation, the bit error probability at the receiver is given by

$$p_b = Q\left(\sqrt{\frac{2E_r^b}{N_0}}\right) \tag{12}$$

where N_0 is the single-sided power spectral density of noise, and the Q function is given by $Q(x) = (1)/(\sqrt{2\pi})\int_x^\infty e^{-x^2/2}\mathrm{d}x$. Suppose that a video packet consists of L bits. In case bit errors are detected, the packet is discarded at the receiver. Then, the packet loss ratio (PLR), denoted by p, is given by $p = 1-(1-p_b)^L$. Let p_k be the probability that the video packet is correctly received after k retransmissions. We have $p_k = p^k(1-p)$. Therefore, the average number of transmissions per packet, denoted by $\bar{N}_t$, is

$$\begin{aligned}\bar{N}_t &= \sum_{k=0}^{N_{\mathrm{MAX}}}(k+1)p_k \\ &= \sum_{k=0}^{N_{\mathrm{MAX}}}(k+1)p^k(1-p) \approx \frac{1}{1-p}, \\ &\quad \text{for sufficiently large } N_{\mathrm{MAX}}\end{aligned} \tag{13}$$

where N_{MAX} is the maximum number of retransmissions and the approximation in (13) is based on the assumption that N_{MAX} is sufficiently large in the WVSN surveillance applications. Combining (11) and (13) yields

$$\bar{N}_t = \frac{1}{\left(1-Q\left(\sqrt{2hd^{-n}E_t^b/N_0}\right)\right)^L}. \tag{14}$$

In this case, the effective transmission energy per bit, denoted by $\bar{E}_t^b$, is given by

$$\bar{E}_t^b = \bar{N}_t E_t^b = \frac{E_t^b}{\left(1-Q\left(\sqrt{2hd^{-n}E_t^b/N_0}\right)\right)^L} \tag{15}$$

and the overall transmission energy is given by $\bar{E}_t^b \cdot R_s$. In Fig. 7, we plot the effective transmission energy per bit $\bar{E}_t^b$ (top) and the average number of transmissions $\bar{N}_t$ (bottom) as functions of E_t^b. The corresponding values of these model parameters are listed in Table I. The simulation result indicates that there is an optimum choice of E_t^b, denoted by $\bar{E}_t^{b,*}$, which minimizes the effective transmission energy per bit $\bar{E}_t^b$ and the overall transmission energy. Intuitively, the existence of $\bar{E}_t^{b,*}$ can be explained by the following observation: if the transmission power is too low, a large number of retransmissions will be requested,

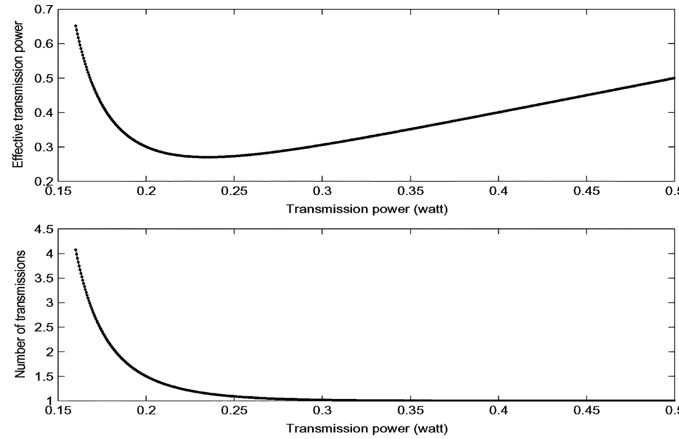

Fig. 7. Effective transmission power per bit.

TABLE I
CONFIGURATION OF THE MODEL PARAMETERS FOR Fig. 7

Parameter	Description	Value
G_t	Transmitter antenna gain	1
G_r	Receiver antenna gain	1
n	Path loss index	2
λ	Wave length	$\frac{1}{9}$ meters
κ	System loss	1
L	Packet size	1024 bits

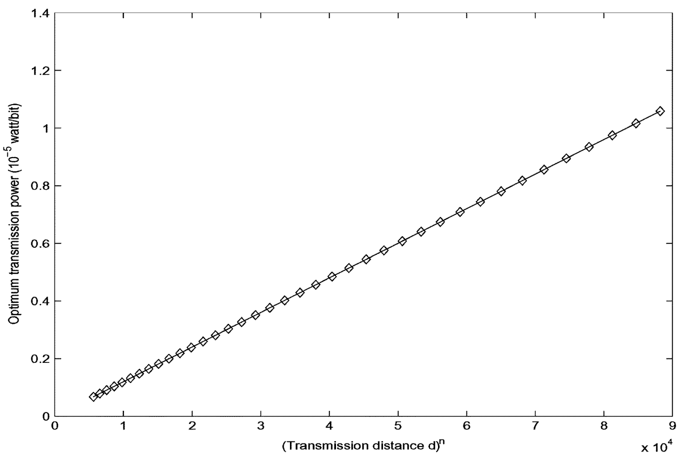

Fig. 8. Minimum transmission power as a function of d^n.

therefore $\bar{E}_t^b$ increases; on the other hand, it is also a waste of energy to use an unnecessarily high transmission power.

Clearly, the optimum transmission energy per bit $\bar{E}_t^{b,*}$ depends on the transmission distance d. The complicated expression in (15) would indicate that $\bar{E}_t^{b,*}$ has a complicated behavior. However, in fact, the relationship between $\bar{E}_t^{b,*}$ and d is quite simple. In Fig. 8, we plot $\bar{E}_t^{b,*}$ as a function of d^n. It can be seen that $\bar{E}_t^{b,*}$ is proportional to d^n, i.e.,

$$\bar{E}_t^{b,*} \approx C_t d^n \tag{16}$$

where C_t is a model constant. For a given transmission distance d, $\bar{E}_t^{b,*}$ minimizes the total transmission energy. Note that, with retransmission, each packet is received correctly without transmission errors, which is the same as the ideal (error-free) case in Section V. Therefore, the AMD analysis developed in Section V can be directly used to analyze the performance limit of a wireless video surveillance sensor by simply replacing the power model in (7) with

$$P_t = \bar{E}_t^{b,*} \cdot R_t = C_t d^n R_t. \tag{17}$$

B. Performance Analysis for Low-Delay Video Monitoring

Within the context of wireless video surveillance, the wireless video sensor is allowed to take advantage of transmission delay and retransmit the video packets if they are not received correctly. However, in WVSN monitoring, because of the stringent delay requirement, packet retransmissions are infeasible even if the packet is received with errors. These transmission errors will be passed to the video decoder and cause decoding errors and error propagation to the subsequent frames. As discussed in Section III, the video distortion caused by the transmission errors is called transmission distortion, denoted by D_t, and the end-to-end distortion is given by $D = D_s + D_t$. Now, the task of wireless video sensors is to perform optimum resource (power) allocation to minimize the end-to-end distortion, as formulated in (3). In this section, based on the P-R-D model and a transmission distortion model developed in our previous work [17], we study the energy consumption of the wireless video sensor and analyze its performance limit.

To study the behavior of transmission errors and analyze their impact on the performance of wireless video sensors, we need a model for transmission distortion. In our previous work [17], we successfully developed a model to predict the picture distortion caused by packet loss. Let p be the packet loss ratio and ξ be the average fraction of intra macroblocks (MBs) in the video frame. The transmission distortion at frame n, denoted by $D_t(n)$, is given by

$$D_t(n) = \Gamma_1 \cdot D_t(n-1) + \Gamma_2 \cdot F_d(n) \tag{18}$$

where $\Gamma_1 = (1-p)(1-\xi)b + p$, and $\Gamma_2 = pa$. Here, a and b are model parameters whose detailed definitions and estimation scheme are given in [17]. $F_d(n)$ is the variance of the motion-compensated original difference picture. Solving the recursive equation (18) yields

$$D_t(n) = \Gamma_1^n \cdot D_t(0) + \Gamma_2 \sum_{i=1}^{n} \Gamma_1^{n-i} \cdot F_d(i). \tag{19}$$

The average transmission distortion of a sequence of video frames, denoted by D_t, is given by [17]

$$D_t = \theta \frac{p}{1-p} \bar{F}_d(n) \quad \text{and} \quad \theta = \frac{a}{(1-b+b\xi)} \tag{20}$$

where $\bar{F}_d(n)$ is the time average of the frame difference $F_d(n)$. (The estimation of the model parameters a and b are explained in [17].) It should be noted that the packet loss ratio p depends on the packet size. Here, we assume that each packet contains a fixed number of MBs. Therefore, the packet size also depends on the actual coding bit rate. The simulation results in [17] demonstrate that the distortion model is very accurate, with a prediction error of less than 5%.

Based on the P-R-D model and the transmission distortion model, we are ready to study the performance limit of the wireless video sensor. Note that

$$D = D_s + D_t = D_s(R_s, P_s) + \theta \frac{p\left(E_t^b\right)}{1 - p\left(E_t^b\right)} \bar{F}_d(n). \tag{21}$$

Here, the packet loss ratio p is a function of the transmission energy per bit E_t^b, denoted by $p(E_t^b)$, whose expression can be

TABLE II
CONFIGURATION OF THE MODEL PARAMETERS IN (22)

Parameter	Description	Value
G_t	Transmitter antenna gain	1
G_r	Receiver antenna gain	1
n	Path loss index	2
λ	Wave length	$\frac{1}{9}$ meters
κ	System loss	1
L	Packet size	512 bits
σ^2	Input video variance	150 (MSE)
θ	See (20)	0.8
$\bar{F}_d(n)$	Average frame difference	200

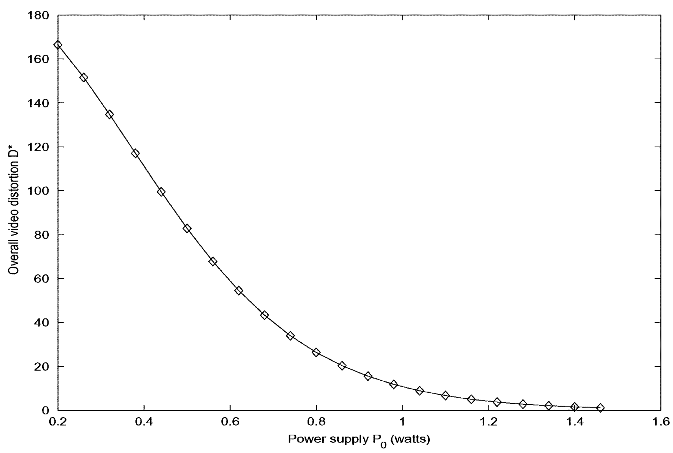

Fig. 9. Minimum video distortion D^* of the wireless video sensor as a function of its power supply P_0.

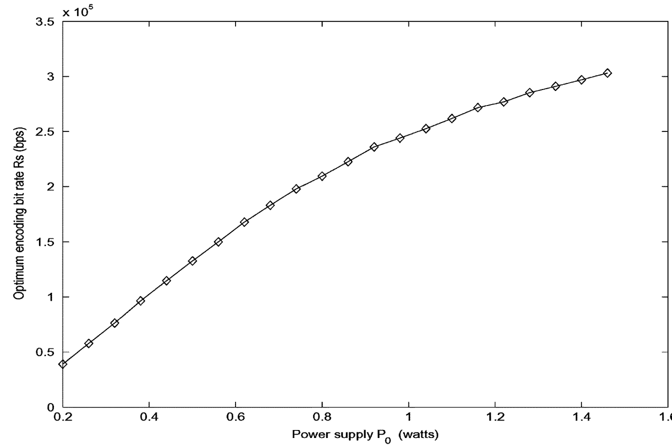

Fig. 10. Optimum encoding bit rate R_s as a function of the power supply P_0.

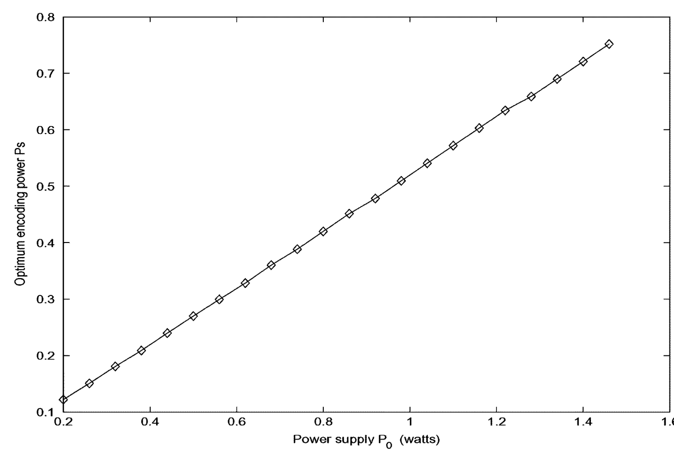

Fig. 11. Optimum encoding power consumption P_s as a function of the power supply P_0.

obtained from (11). To achieve the maximum video sensing performance, the wireless video sensor needs to determine an optimum allocation of its energy resource to minimize the overall video distortion D. This resource allocation problem can be written as

$$\begin{aligned} D^* &= \min_{\{P_s, E_t^b\}} D \\ &= \min_{\{P_s, E_t^b\}} D_s(R_s, P_s) + \theta \frac{p\left(E_t^b\right)}{1 - p\left(E_t^b\right)} \bar{F}_d(n) \\ &\quad \text{s.t. } P_s + E_t^b R_s = P_0. \end{aligned} \tag{22}$$

The solution to (22) represents the optimum strategy for the wireless video sensor to allocate its energy resource, configure its video encoder, and transmit the compressed video data over the wireless channel to the AFN.

To understand how the wireless video sensor achieves its performance limit, as well as the inherent relationship between the performance limit D^* and the physical configuration $\{P_0, d\}$ of the sensor node, we numerically solve the minimization problem in (22). The values of the parameters used for the computation are listed in Table II. The noise power is set to be -65 dBm. In Fig. 9, we plot the minimum distortion D^* as a function of the power supply P_0. Two major observations can be made based on the comparison between Figs. 9 and 6. First, for the same power supply level, the minimum distortion in Fig. 9 is larger. This is because the AMD analysis in Fig. 6 allows a longer delay for packet retransmissions. The second observation is that the minimum distortion in Fig. 9 decays faster when the power supply P_0 increases, especially when the power supply is low. This is because increasing the power supply will significantly improve the channel condition and reduce the transmission distortion. Figs. 10 and 11 show the corresponding encoding bit rate R_s and encoding power consumption P_s as functions of the initial power supply P_0 for the minimum distortion D^* in Fig. 9. As mentioned before, the maximum performance of the wireless video sensor or the minimum video distortion D^* depends on the physical configuration $\{P_0, d\}$ of the sensor node. Therefore, D^* is a function of $\{P_0, d\}$, denoted by $D^*(P_0, d)$. In Fig. 12, we plot the function $D^*(P_0, d)$. Note that when the power supply P_0 is low and the deployment distance d is large, the wireless video sensor has no capability to perform the video compression and transmission task. Even if the sensor node spends some energy to compress the video information, it does not have sufficient energy to send the data to the distant AFN. In this case, no bit is received by the AFN, and the reconstructed video is totally blank. Hence, the corresponding video distortion is equal to the variance of video data. This is why there is a flat region in the upper left corner of the plot. The function $D^*(P_0, d)$ tells us that the minimum video distortion that we can expect

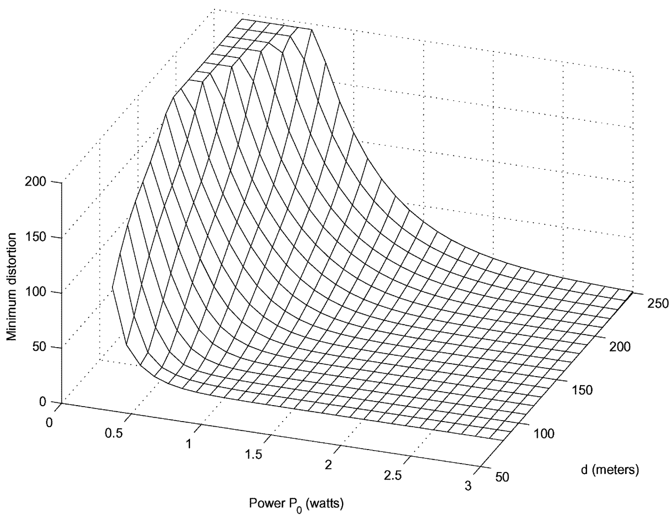

Fig. 12. Minimum video distortion D^* of a wireless video sensor with power supply P_0 and deployment distance d.

from a wireless video sensor if the sensor node is deployed with a power supply level P_0 and a distance from the AFN. Clearly, this function will play a very important role in resource allocation, performance analysis, and topology control of the whole WVSN.

VII. Conclusion and Discussion of Future Work

In this paper, based on our previous study on the P-R-D behavior of an energy-scalable video encoder, we have derived an analytical P-R-D model for video encoding under energy constraints. We then considered two different scenarios for wireless video sensing, studied the energy consumption of wireless video transmission, and analyzed the performance limit of the wireless video sensor for each scenario. The concepts, models, and analysis framework developed in this study give us valuable insights on the complex behaviors of wireless video sensors and provide a theoretical basis for resource allocation and performance analysis in WVSNs.

Acknowledgment

The authors would like to thank the reviewers for their valuable comments.

References

[1] P. Agrawal, J.-C. Chen, S. Kishore, P. Ramanathan, and K. Sivalingam, “Battery power sensitive video processing in wireless networks,” in *Proc. IEEE PIMRC’98*, Boston, MA, Sep. 1998, pp. 116–120.

[2] S. M. Akramullah, I. Ahmad, and M. L. Liou, “Optimization of H.263 video encoding using a single processor computer: Performance tradeoffs and benchmarking,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 11, no. 8, pp. 901–915, Aug. 2001.

[3] I. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, “A survey on sensor networks,” *IEEE Commun. Mag.*, no. 8, pp. 102–114, Aug. 2002.

[4] *AMD PowerNow!TM Technology Platform Design Guide for Embedded Processors* (in AMD Inc.), [Online]. Available: http://www.amd.com/epd/processors

[5] M. Bhardwaj and A. P. Chandrakasan, “Bounding the lifetime of sensor networks via optimal role assignments,” in *Proc. IEEE Infocom*, 2002, pp. 1587–1596.

[6] T. Berger, *Rate Distortion Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1984.

[7] 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 Digit. Computat. Video*, 2001, pp. 8–12.

[8] J. H. Chang and L. Tassiulas, “Energy conserving routing in wireless ad-hoc networks,” in *Proc. IEEE INFOCOM*, Tel Aviv, , Israel, Mar. 2000, pp. 22–31.

[9] L. Chen, Z. He, S. Sethuraman, and C. W. Chen, “MPEG-4 encoder implementation on MAP-CA DSP,” in *Proc. Int. Conf. Consumer Electron.*, Los Angeles, CA, Jun. 2002, pp. 276–277.

[10] T. Chiang and Y.-Q. Zhang, “A new rate control scheme using quadratic rate distortion model,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 7, no. 1, pp. 246–250, Feb. 1997.

[11] P. A. Chou and A. Sehgal, “Rate-distortion optimized receiver-driven streaming over best-effort networks,” in *Proc. Packet Video Workshop*, Pittsburgh, PA, Apr. 2002, pp. 121–130.

[12] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.

[13] W. Ding and B. Liu, “Rate control of MPEG video coding and recording by rate-quantization modeling,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 6, no. 1, pp. 12–20, Feb. 1996.

[14] Y. Eisenberg, C. E. Luna, T. N. Pappas, R. Berry, and A. K. Katsaggelos, “Joint source coding and transmission power management for energy efficient wireless video communications,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 12, Special Issue on Wireless Video, no. 6, pp. 411–424, Jun. 2002.

[15] Z. He and S. K. Mitra, “A unified rate-distortion analysis framework for transform coding,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 11, no. 12, pp. 1221–1236, Dec. 2001.

[16] Z. He and S. K. Mitra, “A linear source model and a unified rate control algorithm for DCT video coding,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 12, no. 11, pp. 970–982, Nov. 2002.

[17] Z. He, J. Cai, and C. W. Chen, “Joint source channel rate-distortion analysis for adaptive mode selection and rate control in wireless video coding,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 12, Special Issue on Wireless Video, no. 6, pp. 511–523, Jun. 2002.

[18] 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.

[19] Z. He, Y. Liang, and I. Ahmad, “Power-rate-distortion analysis for wireless video encoding under energy constraint,” in *Proc. SPIE Visual Commun. Image Process.*, San Jose, CA, Jan. 2004, pp. 57–68.

[20] Z.-L. He, C.-Y. Tsui, K.-K. Chan, and M. Liou, “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.

[21] J. Pan, Y. T. Hou, L. Cai, Y. Shi, and S. X. Shen, “Topology control for wireless video surveillance networks,” in *Proc. ACM Mobicom*, San Diego, CA, Sep. 14–19, 2003, pp. 286–299.

[22] *Wireless LAN medium access control (MAC) and physical layer (PHY) Spec*, IEEE, P802.11/D5, Draft Standard IEEE 802.11, May 1996.

[23] *Intel XScale Technology*, Intel Inc. [Online]. Available: http://www.intel.com/design/intelxscale

[24] N. Jayant, J. Johnston, and R. Safranek, “Signal compression based on models of human perception,” *Proc. IEEE*, vol. 81, no. 10, pp. 1385–1422, Oct. 1993.

[25] C. E. Jones, K. M. Sivalingam, P. Agrawal, and J.-C. Chen, “A survey of energy efficient network protocols for wireless networks,” *Wireless Networks*, vol. 7, pp. 343–358, 2001.

[26] J. M. Kahn, R. H. Katz, and R. S. J. Pister, “Next century challenge: Mobile networking for smart dust,” in *Proc. ACM MobiCom*, Seattle, WA, Aug. 1999, pp. 271–278.

[27] J. Lorch and A. Smith, “Improving dynamic voltage scaling algorithms with PACE,” in *Proc. ACM SIGMETRICS Conf.*, Jun. 2001, pp. 50–61.

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

[29] R. Min, T. Furrer, and A. Chandrakasan, “Dynamic voltage scaling techniques for distributed microsensor networks,” in *Proc. IEEE Computer Soc. Workshop VLSI*, Apr. 2000, pp. 43–46.

[30] A. Ortega and K. Ramchandran, “Rate-distortion methods for image and video compression,” *IEEE Signal Process. Mag.*, vol. 15, no. 6, pp. 23–50, Nov. 1998.

[31] I. M. Pao and M. T. Sun, "Statistical computation of discrete cosine transform in video encoders," *J. Vis. Commun. Image Representation*, vol. 9, no. 2, pp. 163–170, Jun. 1998.
[32] C. E. Perkins, E. M. Belding-Royer, and S. R. Das, "Ad hoc on-demand distance vector (AODV) routing," IETF Internet Draft (in draft-ietf-manet-aodv-09.txt) Nov. 2001.
[33] T. S. Rappaport, *Wireless Communications: Principles and Practice*. Englewood Cliffs, NJ: Prentice-Hall, 1996.
[34] J. Ribas-Corbera and S. Lei, "Rate control in DCT video coding for low-delay communications," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 9, no. 1, pp. 172–185, Feb. 1999.
[35] V. Rodoplu and T. H. Meng, "Minimum energy mobile wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 8, pp. 1333–1334, 1999.
[36] T. Sikora, "The MPEG-4 video standard verification model," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 7, no. 1, pp. 19–31, Feb. 1997.
[37] K. Stuhlmuller, N. Farber, M. Link, and B. Girod, "Analysis of video transmission over lossy channels," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 6, pp. 1012–1032, Jun. 2000.
[38] 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.
[39] T. Wiegand, G. J. Sullivan, G. Bjntegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 560–576, Jul. 2003.
[40] D. Wu, T. Hou, and Y.-Q. Zhang, "Transporting real-time video over the internet: Challenges and approaches," *Proc. IEEE*, vol. 88, no. 12, pp. 1855–1875, Dec. 2000.
[41] F. Ye, H. Luo, J. Cheng, S. Lu, and L. Zhang, "A two-tier data dissemination model for large-scale wireless sensor networks," in *Proc. ACM MobiCom*, Atlanta, GA, Sep. 2002, pp. 148–159.

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

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.

Prof. 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. He was the recipient of the 2002 IEEE Transactions on Circuits and Systems for Video Technology Best Paper Award and the SPIE VCIP Young Investigator Award in 2004.

Dapeng Oliver Wu (S'98–M'04–SM'06) received the B.E. degree in electrical engineering from Huazhong University of Science and Technology, Wuhan, China, in 1990, the M.E. degree in electrical engineering from Beijing University of Posts and Telecommunications, Beijing, China, in 1997, and the Ph.D. degree in electrical and computer engineering from Carnegie Mellon University, Pittsburgh, PA, in 2003.

Since August 2003, he has been with Electrical and Computer Engineering Department, University of Florida, Gainesville, as an Assistant Professor. His research interests are in the areas of networking, communications, multimedia, signal processing, and information and network security. He is an Associate Editor for the *International Journal of Ad Hoc and Ubiquitous Computing*.

Prof. Wu is currently an Associate Editor for the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY and the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY. He is also a Guest Editor for the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS Special Issue on Cross-layer Optimized Wireless Multimedia Communications. He served as Program Chair for IEEE/ACM First International Workshop on Broadband Wireless Services and Applications (BroadWISE 2004); and as a technical program committee member of over 30 conferences. He is Vice Chair of Mobile and wireless multimedia Interest Group (MobIG) of the Technical Committee on Multimedia Communications of the IEEE Communications Society and is a member of the Best Paper Award Committee of Technical Committee on Multimedia Communications of the IEEE Communications Society. He was the recipient of the IEEE Circuits and Systems for Video Technology (CSVT) Transactions Best Paper Award 2001.