Adaptive Sensing and Optimal Power Allocation for Wireless Video Sensors With Sigma-Delta Imager

Mališa Marijan, *Student Member, IEEE*, Ilker Demirkol, *Member, IEEE*, Danijel Maričić, *Student Member, IEEE*, Gaurav Sharma, *Senior Member, IEEE*, and Željko Ignjatović, *Member, IEEE*

Abstract—We consider optimal power allocation for wireless video sensors (WVSs), including the image sensor subsystem in the system analysis. By assigning a power-rate-distortion (P-R-D) characteristic for the image sensor, we build a comprehensive P-R-D optimization framework for WVSs. For a WVS node operating under a power budget, we propose power allocation among the image sensor, compression, and transmission modules, in order to minimize the distortion of the video reconstructed at the receiver. To demonstrate the proposed optimization method, we establish a P-R-D model for an image sensor based upon a pixel level sigma-delta $(\Sigma\Delta)$ image sensor design that allows investigation of the tradeoff between the bit depth of the captured images and spatio-temporal characteristics of the video sequence under the power constraint. The optimization results obtained in this setting confirm that including the image sensor in the system optimization procedure can improve the overall video quality under power constraint and prolong the lifetime of the WVSs. In particular, when the available power budget for a WVS node falls below a threshold, adaptive sensing becomes necessary to ensure that the node communicates useful information about the video content while meeting its power budget.

Index Terms—Image sensors, power-rate-distortion (P-R-D), resource management, sigma-delta $(\Sigma\Delta)$ modulation, wireless video.

I. INTRODUCTION

IN typical video communication systems, one of the major problems encountered in system optimization is the control of the system performance under bandwidth constraints. To analyze the behavior of the system under bandwidth constraints, the rate-distortion (R-D) theory is usually applied. The R-D characteristics of a system are used to solve the problem of finding the minimum number of bits to be transmitted to achieve a given level of distortion.

A wireless video sensor network is a system that contains spatially distributed wireless video sensors (WVSs). The function of the WVSs (sensor nodes) is to capture visual information about the environment, to compress the sensed data and to transmit the compressed data through the wireless medium. Three major modules of a WVS are therefore image sensing, video compression, and wireless transmission.

However, wireless video sensors usually operate under limited power supply, where the available power plays an important role in the resulting video quality as well as in the life-time of the system. Therefore, the power allocation among the image sensor, compression, and transmission modules in the sensor nodes is an important aspect of system optimization. Since wireless networks are limited in both power and bandwidth, rate-distortion analysis is an insufficient tool for optimal resource allocation.

This problem of the deficiency of the R-D analysis for wireless video sensor networks has been addressed in [1]. It has been suggested that for optimal resource allocation, classical R-D analysis must be extended to include additional resource constraints. A new, power-rate-distortion (P-R-D) analysis has to be applied for power and bit allocation in wireless video systems, where the two different concepts, power consumption and R-D analysis, must be merged together [1], [2].

In wireless local area networks (WLAN), which typically operate in a 50–100 m communication range, a significant portion of the total power of a WVS is consumed by the compression and transmission modules [3]. To maximize video quality under power and rate constraints, a P-R-D model for the video compression module of a WVS is first developed in [3]. The optimization framework based upon the P-R-D model of the video encoder is further used in [2] to analyze the power tradeoff between the video encoding and wireless data transmission modules. However, simulation results from [4] indicate that in a 10–20 m communication range, the camera consumes almost 50% more power than wireless transmission. Still, the image sensor subsystem of the WVS has not been included in the optimization procedure. One of the main reasons for excluding the imager from the system optimization is that it is not well understood how to incorporate image sensor characteristics within the existing video encoder optimization framework [4]. In addition, the power-quality tradeoff of an image sensor depends strongly upon the specific sensor design, which makes attempts to derive general specifications rather difficult.

In a wireless personal area network (WPAN), communication is typically performed within a 10 m range. For the same data rate as in WLAN, the power requirements for the transmission module can be significantly reduced, resulting in a situation where the image sensor has a higher impact on the overall power consumption of a WVS. Therefore, resource allocation in WVS for WPAN cannot be performed in an optimal way

without considering the image sensor in the system analysis. If the sensor is included in the power control, the power consumption and rate of the sensor can be optimally set and reduced as the power budget decreases. Lowering the power consumption of the image sensor when the power budget decreases, can increase overall video quality because more power can be allocated to the compression and transmission modules.

In addition, power-sensitive multimedia applications, such as 3 G phones, drive the market towards low power solutions for video encoders. For low power performance, real-time MPEG-4 encoders employ dedicated hardware accelerators [7]. For example, according to results from [8], a hardware optimized MPEG-4 encoder in 60 nm technology can deliver high quality images (VGA at 30 fps) consuming less than 5 mW. Also, although the CMOS image sensor technology scales at the same pace as the standard digital CMOS process, it is lagging behind the standard digital CMOS process by several technology nodes [9]. The reason for this lag is that the smaller pixel size has a negative effect on both the electrical and optical performance of the imager so that any further scaling of the pixel size has proven rather challenging. In contrast to the video encoders implemented in digital CMOS technologies, conventional analog CMOS image sensors cannot take full advantage of power reduction offered by current CMOS technology progress.

On the other hand, several emerging fully-digital CMOS image sensor designs [10] based upon pixel-level sigma-delta ($\Sigma\Delta$) A/D conversion methods, denoted as $\Sigma\Delta$ imagers, with low power feature and unique power quality tradeoff have been reported [11]–[13]. The technology predominantly targets surveillance imaging applications where high dynamic range, good low-light response, and low power are in demand. In addition, an oversampling technique exploited in the $\Sigma\Delta$ conversion allows for a simple tradeoff between the frame rate and the bit resolution [11], [12], and offers relatively simple power control [13]. This controllability of the design enables the $\Sigma\Delta$ imager to meet many application-specific needs. In particular, a low power consumption and design controllability make the $\Sigma\Delta$ imager suitable for WVS with adjustable parameters.

In this paper, we introduce a P-R-D concept for an image sensor which then can be used to build a comprehensive P-R-D framework for WVSs optimization. This approach allows us to establish an optimal solution for the power allocation problem for WVS containing an arbitrary image sensor with variable output rates. We further consider $\Sigma\Delta$ image sensor design, for which power, output bit rate, and video quality can be controlled with oversampling. Based upon dependence of the power, rate, and distortion with the oversampling, we derive a P-R-D model of a $\Sigma\Delta$ imager. The sensor model is then used within developed optimization framework to investigate the tradeoff between the bit depth of the captured images and spatio-temporal characteristics of the video sequence under power constraint.

The rest of the paper is organized as follows. The optimization problem for WVS containing the image sensor with variable output rate is consider in Section II. In Section III, an analytic P-R-D model for a $\Sigma\Delta$ imager is proposed and evaluated with the experimental results from the $\Sigma\Delta$ imager designed in our lab. Using developed optimization framework and model
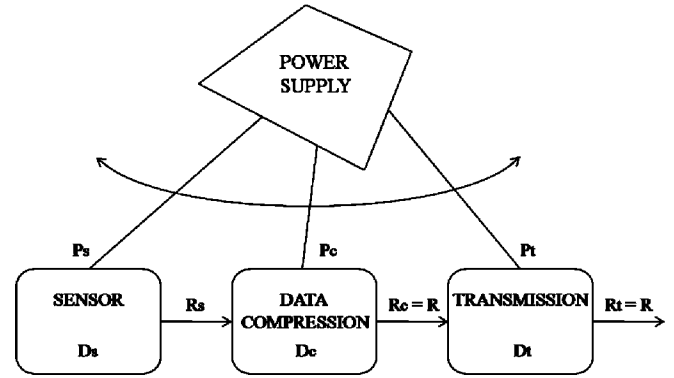

Fig. 1. Wireless video sensor: power has to be optimally allocated among the image sensor, compression, and transmission modules.

for the $\Sigma\Delta$ imager, P-R-D analysis for a WVS is carried out in Section IV. Our work is summarized in Section V.

II. P-R-D OPTIMIZATION FRAMEWORK

In this section, the problem of optimal power allocation among the image sensor, compression, and transmission modules in a WVS is analyzed. The P-R-D optimization framework for the system is derived under the assumption that the image sensor is capable of adjusting its output rate according to the available power budget.

A simple model of a WVS is shown in Fig. 1. The model consists of three modules: image sensor, video compression, and transmission. In our framework, we assume that the relationship between the power, rate, and distortion of each module is well described. Power-rate-distortion functions are denoted as $P_s - R_s - D_s$, $P_c - R_c - D_c$, and $P_t - R_t - D_t$, for the sensor, compression, and transmission, respectively. We also assume that the total available power provided to a WVS and the target bit rate are known and equal to P and R, respectively. The optimization problem can be stated as: *How should power P be allocated among the image sensor, compression, and transmission modules to minimize the overall distortion introduced for a given target rate R?*

The most common measure for video quality is the end-to-end mean squared error (MSE) between the original and the received picture frames, referred here as the distortion D [2]. Distortion in each module can be defined as the mean squared difference between the output and input frame of the module. In [5], it is shown that compression and transmission distortions are uncorrelated, so that the total distortion of the two subsystems becomes the sum of the individual distortions. Herein, we treat the overall distortion of a WVS as the sum of the distortions over all three subsystems. In addition, any dependence that might exist between the sensor and the compression distortions are reflected in the dependence of the compression module distortion on the sensor output rate.

We consider a special case where the image sensor output rate R_s can be represented as a function of the power consumption P_s (as in the case of the $\Sigma\Delta$ imager discussed in Section III), so that the image sensor distortion D_s depends only upon the power P_s. We also assume that the bit rates R_c and R_t are equal
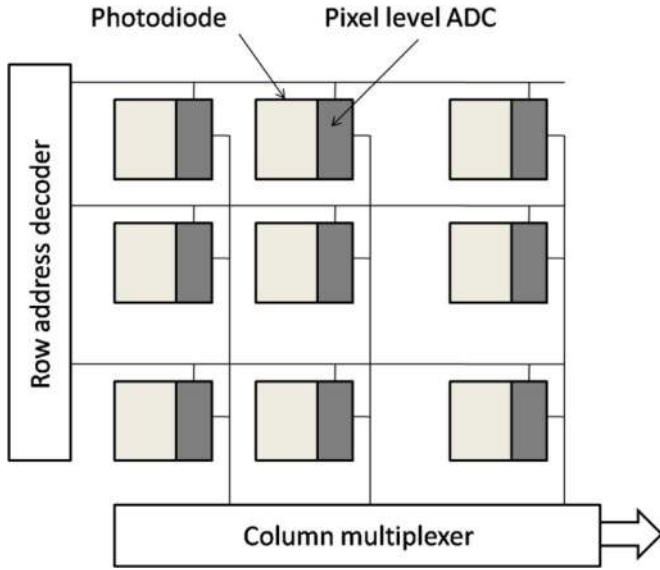
Fig. 2. DPS architecture: analog-to-digital conversion is performed at every pixel location. Using the row decoder and the column multiplexer, pixel bits are read out in a fashion similar to digital memory readout [10].
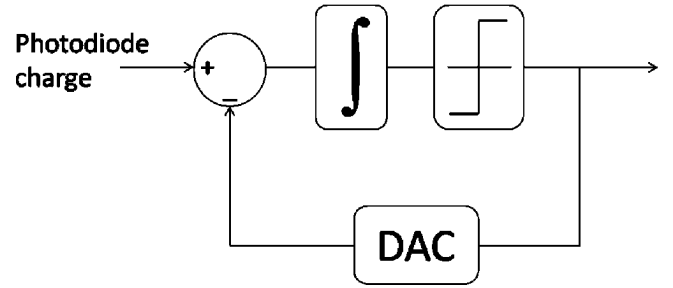

Fig. 3. $\Sigma\Delta$ imager: the charge from each photodiode is read by a $\Sigma\Delta$ modulator consisting of an integrator, a single bit comparator, and a single bit feedback digital-to-analog converter.

to the target bit rate R (Fig. 1). Hence, for a given rate R, all three distortion functions can be described as a function of the power allocated to the corresponding module. In addition, the distortion D_c that characterizes the compression module is also a function of the input rate R_s and, hence, depends upon the image sensor power consumption.

The optimization problem can be mathematically formulated as shown in (1)–(2) at the bottom of the page. The solution to this problem provides the optimal power allocation for the image sensor, compression, and transmission modules such that overall distortion is minimized. In addition, for each power level P, the optimization procedure can be repeated to find an optimum rate R that will provide minimum distortion. If the system is designed to support different target bit rates, controlling the output rate of the WVS can lead to improved video quality, as shown in [2].

The optimal solution can be found using the Lagrange multiplier method, which formulates the optimization problem as a problem of finding the minimum of the unconstrained function $\wedge = \wedge(P_s, P_c, P_t, \lambda)$ defined as

$$\wedge(P_s, P_c, P_t, \lambda) = D_s(P_s) + D_c(R_s(P_s), P_c) + D_t(P_t) \\ + \lambda(P_s + P_c + P_t - P) \quad (3)$$

where λ is a new variable called the Lagrange multiplier. If we assume that the distortion functions are differentiable functions,

the necessary conditions for the solution to be optimal can be calculated from the system of equations

$$\frac{\partial D_s}{\partial P_s} + \frac{\partial R_s}{\partial P_s}\frac{\partial D_c}{\partial R_s} = \frac{\partial D_c}{\partial P_c} = \frac{\partial D_t}{\partial P_t} \quad (4)$$

$$P_s + P_c + P_t = P. \quad (5)$$

From (4), it can be seen that, if the output rate of the sensor is fixed, at the optimal solution P_s^*, P_c^*, P_t^*, the tangents of the distortion functions D_s, D_c, and D_t must have the same slope. However, if the output sensor rate can vary with the power consumption, the solution for the power allocation problem for which the distortion is minimized defines the bit depth of the captured images.

III. P-R-D MODEL FOR $\Sigma\Delta$ IMAGE SENSOR

In this section, a P-R-D model for a specific image sensor is derived. A digital pixel sensor (DPS) architecture is considered, where analog-to-digital conversion is performed simultaneously at every pixel location (Fig. 2). Due to a high degree of parallelism, the requirement for conversion speed is relaxed, which translates into low power consumption [10], an important feature for applications with limited power budget, such as WVSs.

A $\Sigma\Delta$ imager is a DPS architecture where the $\Sigma\Delta$ conversion is performed at each pixel site. During the exposure time of the imager, a $\Sigma\Delta$ modulator reads charge from the photodiode *OSR* (oversampling ratio) number of times, and each value characterizes with one bit (see Fig. 3). A decimation filter is then applied to convert the oversampled single-bit stream to a multibit sample at the frame rate.

The $\Sigma\Delta$ imager allows for relatively simple P-R-D modeling. Power, rate, and distortion performance of the imager can be controlled by changing the number of single-bit pixel values that are read every second. In other words, the parameter *OSR* can

$$\min_{P_s, P_c, P_t} \quad D(P_s, P_c, P_t) = D_s(P_s) + D_c(R_s(P_s), P_c) + D_t(P_t) \quad (1)$$

$$\text{subject to} \quad P_s + P_c + P_t = P. \quad (2)$$

be used as a control parameter for the P-R-D modeling and resource allocation, while keeping other design parameters such as power supply voltage fixed. A parametric analysis, denoted as *OSR*-domain analysis, is applied to study resource tradeoffs in the $\Sigma\Delta$ image sensor. Rate-oversampling (R-OSR), distortion-oversampling (D-OSR), and power-oversampling (P-OSR) functions can be derived for the $\Sigma\Delta$ imager. By eliminating the parameter *OSR* from the models, a unique P-R-D relationship can be established.

To characterize the distortion of the $\Sigma\Delta$ image sensor, we assume that the sensor is equipped with ideal optics and that the distortions come only from the imager electronics noise sources. The noise sources and their dependence upon the control parameter *OSR* are discussed in Section III-A. The relevant characteristics for the $\Sigma\Delta$ imager are derived in Section III-B. The experimental results are obtained in Section III-C to evaluate the proposed model considering practical parameters.

A. Noise Model

The noise sources in image sensors can be divided into temporal and spatial [9], [14]. Temporal noise varies in time but is independent across the sensor array; spatial noise represents variations among the pixel values under uniform illumination [15].

The temporal noise sources in the image sensor are the major limiting factor for image sensor performance, particularly under low light conditions and in video applications [16]. Therefore, we simplify the derivation of the noise model by neglecting the spatial component of the noise. In this fashion, the noise model and following P-R-D model can be derived considering the behavior of a single pixel only.

In Section II, it is shown that the optimal solution to the power allocation problem depends upon the rate of change of the distortion function with respect to the power, or equivalently with respect to our control parameter *OSR*. To emphasize the dynamic of the noise power with respect to the parameter *OSR*, we classify the temporal noise sources into three categories: quantization noise, pixel reset and thermal noise, and noise invariant to the oversampling.

1) Quantization Noise: In order to simplify the calculation of the quantization noise, we consider a linearized model of the $\Sigma\Delta$ modulator. In addition, we assume that the decimation filter is an ideal low pass filter. Quantization noise injected by a 1-b comparator is assumed to be a stationary zero-mean white random process with uniform distribution and known variance σ_q^2. The quantization noise power is given as $\sigma_q^2 = \Delta^2/12$, where Δ is the maximum voltage swing of the quantizer proportional to the power supply voltage. Also, we assume that the quantization noise is independent from the input signal. Due to the oversampling, quantization noise occupies OSR times larger bandwidth than the input signal. In addition to the oversampling, the $\Sigma\Delta$ modulator shapes the quantization noise toward high frequencies with a high-pass filter $H(z) = 1 - z^{-1}$, as illustrated in Fig. 4. A digital decimation filter is then used to remove the noise outside the signal band [17]. The quantization
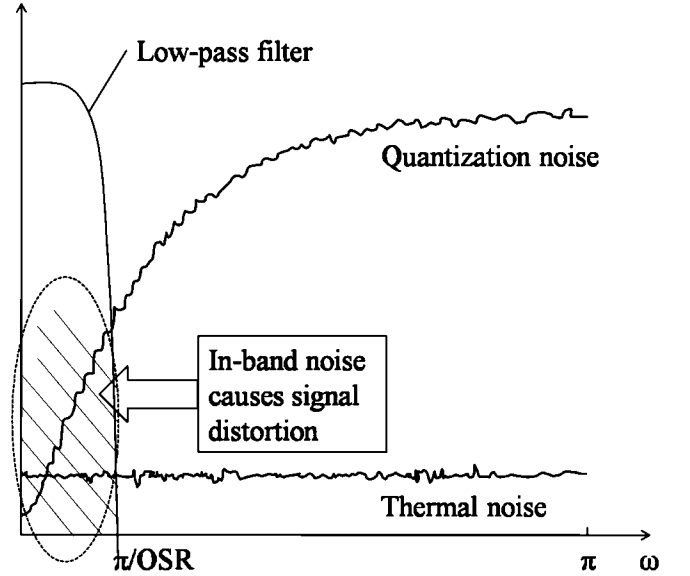

Fig. 4. Oversampling and noise shaping of $\Sigma\Delta$ modulator: quantization noise is shaped outside the signal band and then removed by a decimation filter. However, the termal noise is filtered out without shaping.

noise that remains within the signal band causes the signal distortion. The resulting in-band quantization noise power can be calculated as

$$
\begin{aligned}
\sigma_{q,\Sigma\Delta}^2 &= \int_{-\pi/OSR}^{\pi/OSR} |1 - e^{-jw}|^2 \cdot \frac{\sigma_q^2}{2\pi} dw \\
&= \frac{2\sigma_q^2}{\pi} \left(\frac{\pi}{OSR} - \sin\frac{\pi}{OSR} \right) \\
&\approx \frac{\pi^2 \sigma_q^2}{3 OSR^3}
\end{aligned}
\tag{6}
$$

where the approximation is obtained from the Taylor series expansion for $OSR \gg \pi$. In addition, since the output from the $\Sigma\Delta$ modulator is decimated with a relatively long decimation filter, it can be assumed that the resulting noise after the decimation operation has a Gaussian distribution with power equal to $\sigma_{q,\Sigma\Delta}^2$.

B. Pixel Reset and Thermal Noise

The thermal reset noise is injected into the photodetector each time the photodetector is reset to some initial state after every exposure [18] or it can be injected by resetting the transistor in the feedback of the $\Sigma\Delta$ modulator [13]. In addition, the thermal noise from the pixel transistors contributes to the overall thermal noise. We assume that the overall thermal noise can be represented as a stationary zero-mean white Gaussian process with variance σ_t^2. The thermal noise appears at the input of the $\Sigma\Delta$ modulator and, as opposed to the quantization noise, it is not shaped to a high frequency region (Fig. 4). Since the thermal noise is not shaped, its power after the decimation operation is reduced only by the oversampling ratio OSR, i.e.

$$
\sigma_{t,OSR}^2 = \int_{-\pi/OSR}^{\pi/OSR} \frac{\sigma_t^2}{2\pi} dw = \frac{\sigma_t^2}{OSR}.
\tag{7}
$$

1) Noise Invariant to Oversampling: We consider the shot noise to be a dominant noise source that does not depend upon the control parameter *OSR*. The photodiode shot noise has two independent components: *photon shot noise* caused by the temporal variation in the photon-generated current due to the random arrival nature of the incident photons, and *dark current related shot noise* caused by the variations in the photodiode dark current (i.e., variations in the current generated when no illumination is present). Shot noise has a Poisson distribution and its variance σ_{shot}^2 depends upon the light intensity, dark current value, and exposure time [9], [15], [19]. For the fixed exposure time and photon count, the shot noise power remains unchanged regardless of the number of times charge is read from the photodiode (i.e., the shot noise power is invariant to the parameter *OSR*).

The quantization, thermal, and shot noise are assumed to be independent; therefore, the total noise power σ_N^2 is the sum of the individual components, which is described by

$$
\begin{aligned}
\sigma_N^2 &= \sigma_{q,\Sigma\Delta}^2 + \sigma_{t,OSR}^2 + \sigma_{\text{shot}}^2 \\
&= \frac{\pi^2 \sigma_q^2}{3 OSR^3} + \frac{\sigma_t^2}{OSR} + \sigma_{\text{shot}}^2 \\
&= \sigma_q^2 \left(\frac{\pi^2}{3 OSR^3} + \frac{1}{OSR} \frac{\sigma_t^2}{\sigma_q^2} + \frac{\sigma_{\text{shot}}^2}{\sigma_q^2} \right) \\
&= \sigma_q^2 \left(\frac{\pi^2}{3 OSR^3} + \frac{\alpha}{OSR} + \beta \right)
\end{aligned}
\tag{8}
$$

where the parameters α and β quantify the fraction of the thermal and shot noise power with respect to the quantization noise power. The parameters α and β depend upon the specific $\Sigma\Delta$ image sensor design and can be experimentally determined. In addition, the parameter β depends upon the incident illumination since the photon shot noise depends upon the incident photon count.

C. $\Sigma\Delta$ Image Sensor Characteristics

1) Rate and Distortion Characteristics: The noise model (8) characterizes the distortion for the $\Sigma\Delta$ imager. However, it is convenient to represent the model of (8) in terms of MSE in order to have a representation that is identical to the representation for the compression and the transmission modules. Thus, the distortion oversampling characteristic for the $\Sigma\Delta$ imager is given by

$$
D_s(OSR) = \sigma_q^2 \left(\frac{\pi^2}{3 OSR^3} + \frac{\alpha}{OSR} + \beta \right)
\tag{9}
$$

where in $\sigma_q^2 = \Delta^2/12$, Δ represents the maximum quantization step size in MSE, e.g., $\Delta = 255$ for 8-b representation. The distortion defined by the parameter β characterizes the asymptotic performance of the $\Sigma\Delta$ imager with respect to the parameter *OSR*. The shot noise sets the fundamental limit on the maximum rate and defines the minimum achievable distortion $\beta\sigma_q^2$.

The rate of the image sensor R_s represents the average number of bits to be allocated to each pixel to achieve a distortion of D_s. To calculate the output sensor rate, we assume that the pixel values can be characterized as zero mean random
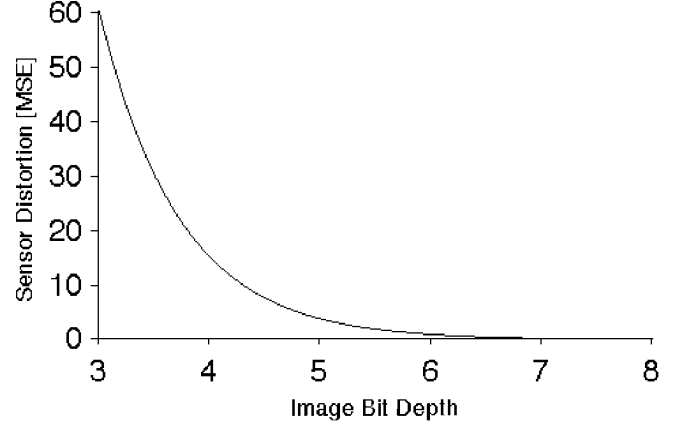

Fig. 5. Rate-distortion characteristic for the sigma-delta ($\Sigma\Delta$) image sensor based upon "Foreman" video sequence.

variables drawn from stationary and ergodic random process. Let $X(i,n)$ be the true photodetector value of a pixel defined by the index i in the nth video frame, and $\hat{X}(i,n)$ be the value at the output of the corresponding pixel. The average image sensor distortion represents a mean squared error between $X(i,n)$ and $\hat{X}(i,n)$, i.e.,

$$
D_s = E\{(X(i,n) - \hat{X}(i,n))^2\}
\tag{10}
$$

where the expected value of variable $x(i,n)$, $E\{x(i,n)\}$, is calculated by averaging values of $x(i,n)$ across the sensor array and across all video frames. We assume that the image sensor distortion has a Gaussian distribution since it contains multiple independent noises filtered by the decimation filter. We also assume that the total distortion can be treated as uncorrelated with the input signal. The image sensor bit depth R_s is the number of bits representing each pixel value regardless of the correlation among the signal values that might exist from frame to frame or from pixel to pixel (i.e., temporal and spatial redundancy). In other words, the image sensor treats the images as if they are generated from a stationary memory-less source. If we also assume that the pixel values are independent and identically distributed Gaussian random variables [20], the R-D characteristic for the sensor can be obtained in the form

$$
D_s(R_s) = \sigma_s^2 2^{-2R_s}
\tag{11}
$$

where $\sigma_s^2 = E\{X(i,n)^2\}$ represents the variance of the video sequence. As an example of the R-D characteristic for the image sensor, we show the characteristic of the sensor that captures the "Foreman" QCIF video sequence at 15 fps. Since the standard video sequences are already quantized, we first relate σ^2 to the variance of the quantized test sequence $\tilde{\sigma}^2$ using Sheppard's corrections [21]

$$
\sigma_s^2 = \tilde{\sigma}^2 - \frac{\Delta^2}{12 \cdot 2^{2R_s}}.
\tag{12}
$$

However, since the test sequences have $R_s = 8$ b, we can assume $\sigma_s^2 \approx \tilde{\sigma}^2$. We estimate $\sigma_s^2 = 3926$ MSE, and the R-D characteristic is shown in Fig. 5.

If we define the statistical parameter γ as $\gamma = \sigma_s^2/\sigma_q^2$, combining (9) and (11) we obtain the relationship between the bit rate R_s and the control parameter *OSR*

$$R_s(OSR) = \frac{1}{2}\log_2\left(\frac{3\gamma OSR^3}{\pi^2 + 3OSR^2(\alpha + \beta \cdot OSR)}\right). \quad (13)$$

2) Power-Quality Tradeoff: Power consumption of the imager contains two components: static and dynamic. We assume that the static power consumption can be neglected with respect to the dynamic power. Since the dynamic power consumption in CMOS circuits is observed to be a linear function of the sampling frequency, the following relation can be obtained for the power as a function of the OSR

$$P_s(OSR) = \frac{OSR}{OSR_{\max}}P_{s\,\max} \quad (14)$$

where $P_{s\,\max}$ is the image sensor power consumption at the maximum output rate $R_{s\,\max}$, i.e., at the maximum oversampling ratio $OSR_{\max}$. We can simplify the notation by denoting $P_{s,n}$ as the normalized sensor power with respect to the maximum power consumption $P_{s\,\max}$. Combining (9) and (14), we obtain a P-D characteristic for the $\Sigma\Delta$ image sensor

$$D_s = \sigma_q^2\left(\frac{\pi^2}{3OSR_{\max}^3}P_{s,n}^{-3} + \frac{\alpha}{OSR_{\max}}P_{s,n}^{-1} + \beta\right). \quad (15)$$

The advantage of the $\Sigma\Delta$ imager in comparison with other image sensor designs is that, by changing the OSR, it allows adjustment to a desired power budget with controlled scaling of the distortion. The noise sources that increase as the available power decreases are quantization and thermal noise. The quantization noise is the major noise source that determines the dynamics of the power allocation, and determines the distortion at low power levels because of its cubic dependence upon the power consumption.

The output rate of the sensor can be related to its power consumption by combining (13) and (14)

$$R_s = \frac{1}{2}\log_2\frac{3\gamma OSR_{\max}^3 P_{s,n}^3}{\pi^2 + 3OSR_{\max}^2 P_{s,n}^2(\alpha + \beta OSR_{\max}P_{s,n})}. \quad (16)$$

Since the sensor rate can be expressed as a function of the sensor power consumption, the power allocation problem can be solved only considering the power allocation for each module, as discussed in Section II. Equation (16) illustrates how sensor rate is determined by the optimal solution to the power allocation problem (P_s^*). The available power budget determines the bit depth of the captured images so that the overall system distortion is minimized.

D. Experimental Results From $\Sigma\Delta$ Image Sensor

To obtain performance parameter values from the $\Sigma\Delta$ imager, we use the $\Sigma\Delta$ imager developed in our laboratory, which is described in [13]. In this design, each photodiode performs the integration function for $\Sigma\Delta$ modulation. A 1-b quantizer is shared among pixels of the row, and a three transistor DAC is used to feed back a fixed amount of charge to the photodiode

when the quantizer is triggered. The focal plane area of the imager is 128×128 pixels with pixel pitch of 10 μm. The $\Sigma\Delta$ imager has been fabricated in 0.35 μm CMOS technology with a 3.3 V power supply.

1) Test Setup: The output bit stream from the $\Sigma\Delta$ imager is decimated with an external FPGA. The single bit stream outputs from the $\Sigma\Delta$ modulators are read from the sensor through the 8-pin port. A 256×12-b ROM is used to store filter coefficients, and 1 K $\times$ 23-b ROM stores control signals that are output from the FPGA to the imager. The imager is equipped with an S-mount 1.7 f micro lens. A VGA component on the FPGA reads the 16 K $\times$ 12-b frame buffer and outputs it to the standard VGA monitor, enabling a real time display of the images.

2) Results: First, the power consumption of the analog part of the sensor is measured. At the frame rate of 30 fps and oversampling ratio of 256, the average power per pixel is estimated to be 16 nW.

We further evaluate the parameters in the noise model (8). For the 1-b quantizer, the quantization step size is equal to the maximum voltage swing of the quantizer. The output voltage swing is measured to be $\Delta = 66$ mV, which gives $\sigma_q^2 = 363$ mV^2. Pixel noise is measured under dark conditions, and the noise power as a function of the oversampling ratio is shown in Fig. 6(a). To illustrate the accuracy of the proposed model, we sweep the *OSR* up to the values of 1500. Two curves that describe the noise model (8) are plotted for the model parameters $\alpha = 0$ and $\alpha = 4.5 \times 10^{-5}$. As predicted by the model, the quantization noise is dominant at low *OSR* values and decays much faster with increased *OSR* than the thermal noise component. The thermal noise becomes dominant at high *OSR* values and must be taken into account.

Since our sensor is optimized for the maximum $OSR_{\max}$ of 256, the optimal value for parameter α was chosen for this range to better fit our model to the experimental results, as shown in Fig. 6(b). The value of 1.7×10^{-4} is obtained by curve fitting, which provides 5–6 dB improvement for $OSR_{\max}$ over the case when the thermal noise is not considered ($\alpha = 0$). The slight deviation of the model from the experimental results for lower OSR values is caused by the transistor flicker (low frequency) noise not included in our model. On the other hand, when the experiment is repeated under normal light conditions, the shot noise level was below the quantization and thermal noise floor. Therefore, we neglect parameter β parameter in further analysis.

Sample images from the $\Sigma\Delta$ imager are captured for different power levels and shown in the Fig. 7. The maximum power level corresponds to an oversampling ratio of 256. The degradation of the image quality becomes apparent at power levels that are about ten times lower than the maximum power consumption. This illustrates that the $\Sigma\Delta$ image sensor design has a very efficient power-quality tradeoff and is suitable for control of WVS performance under power constraint.

IV. P-R-D PERFORMANCE ANALYSIS

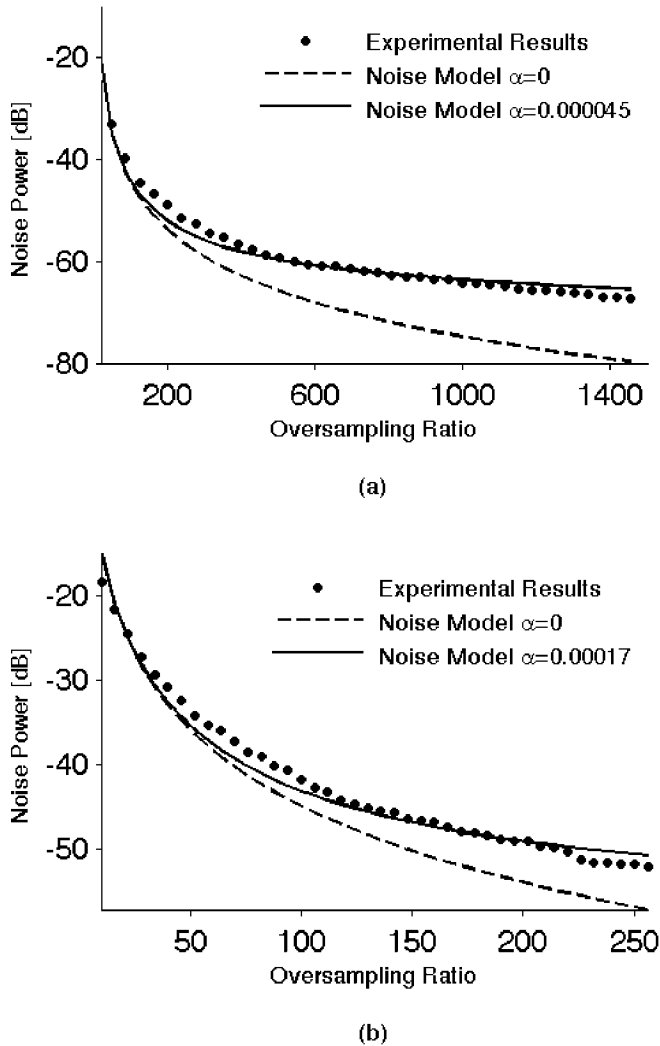In this section, we consider concrete instantiation of the P-R-D optimization framework for a WVS, using the P-R-D

(a)

(b)

Fig. 6. Noise model for $\Sigma\Delta$ imager in dB scale: (a) $OSR\max = 1500$,(b) $OSR\max = 256$. The parameter α accounts for the thermal noise for higher oversampling values.
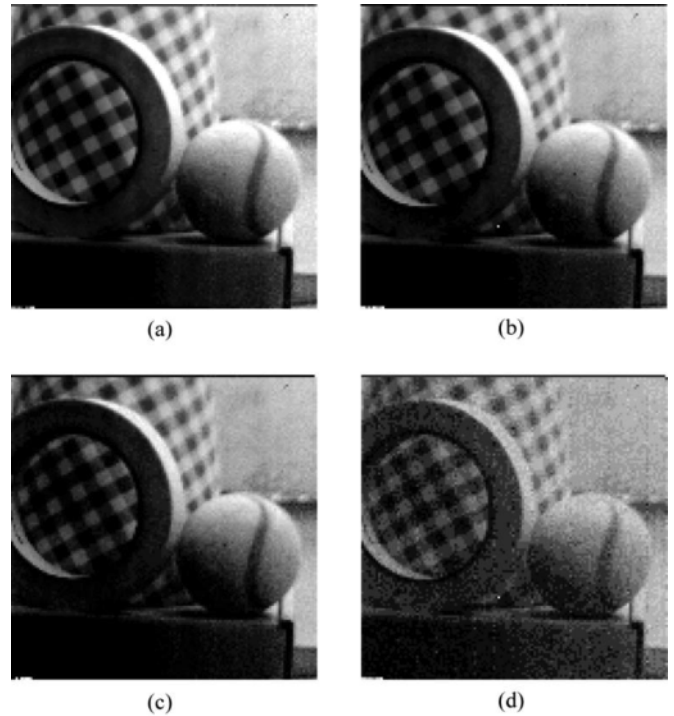

Fig. 7. Images captured by the $\Sigma\Delta$ image sensor: images taken at 100%, 25%, 12.5%, and 6.25% of the maximum power consumption are shown in (a), (b), (c), and (d) respectively.

model of the $\Sigma\Delta$ imager. In Section IV-A, we describe the P-R-D behavior of the individual modules of WVSs. In Section IV-B, we present the results of the optimization analysis.

A. System Description

In order to investigate the power allocation within WVSs, we consider an example of a WVS network. We assume that WVSs are deployed in a small area with a radius of 20 m. The maximum available power budget for the sensor node is 6 mW. Each of the WVSs is supplied with the $\Sigma\Delta$ imager designed to capture grayscale images in QCIF (176×144) format at 15 fps. When the system is first deployed, the image sensor is set to capture images at 8-b resolution, i.e., $OSR_{\max} \approx 75$. The video sequence is then compressed to reduce the bandwidth requirements, and the compressed data is transmitted over the wireless medium. We describe the individual modules:

1) $\Sigma\Delta$ Image Sensor: We assume that the power-quality tradeoff of the $\Sigma\Delta$ image sensor is described by (15) and (16). We chose the parameter values $\alpha = 1.7 \times 10^{-4}$ and $\beta = 0$ to

quantify the influence of the sensor noise. Since the digital logic of our system is implemented externally on the FPGA, the total power consumption of the imager including all necessary processing and control functionality integrated on a single die could not be measured from the tested system and must be estimated. In the $\Sigma\Delta$ design [12], it is shown that the decimation filters fabricated in 0.18 ìm technology consume 560 nW per pixel at $OSR = 1000$ and 50 fps. We assume the similar scenario for the decimation stage of our imager. However, the power value must be scaled to account for the technology difference, as well as for the different oversampling value and frame rate. We assume 50% reduction in the power consumption with new technology node [22], and we estimate the power consumption of the $\Sigma\Delta$ imager to be about 53 nW per pixel, or $P_{s\max} = 1.35$ mW.

2) Video Encoder: In the description of the optimization framework (Section II), we have seen that the compression distortion is a function of the image sensor output rate. For each input rate, there is an optimal design of the video encoder. Here, we briefly describe the performance of the standard compression module optimally designed to encode 8-b input, when images with different bit depths are applied.

We perform an experiment where the standard QCIF video sequences at 15 fps are corrupted with a white Gaussian noise and then passed through the standard MPEG-4 encoder. The compression distortion as a function of image bit depth R_s for different values of the parameter R_c is shown in Fig. 8(a) for the "Coastguard" video sequence. It can be seen that for a sufficiently large bit depth (e.g., $R_s > 4$ b), the compression distortion exhibits very low dependence upon the input rate. We further plot the compression distortion as a function of R_c for different values of the parameter R_c in Fig. 8(b). We can verify that
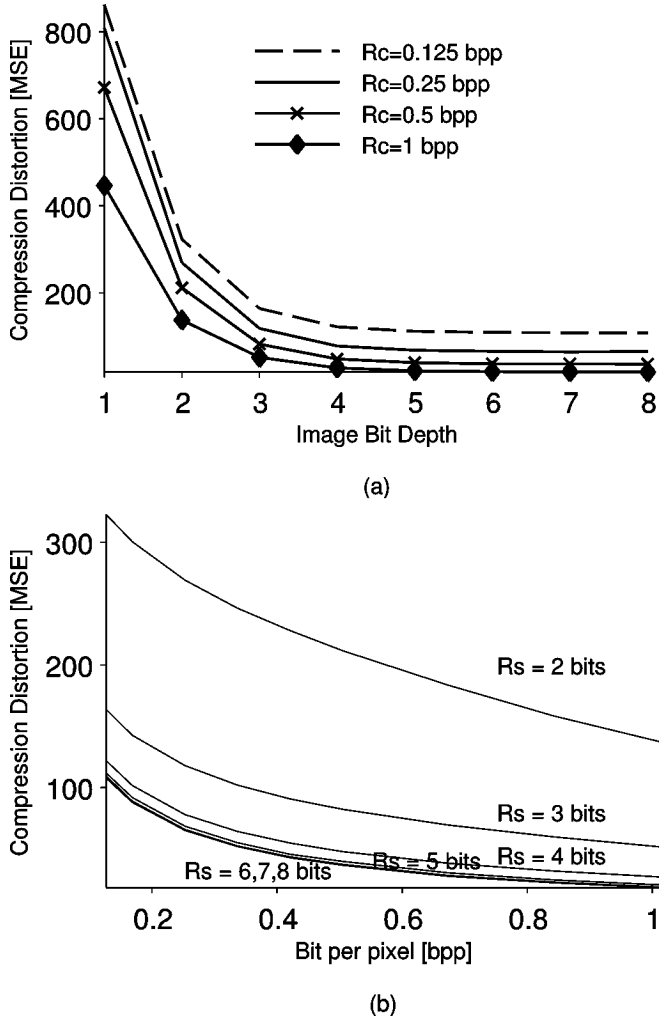
Fig. 8. Compression distortion for the "Coastguard" video sequence: (a) as a function of the image bit depth for different target rates, (b) as a function of the target bit rate for different values of the image bit depth.

there is a very small difference between the distortion curves for higher bit depth. Similar results are obtained for the other video sequences.

Therefore, if the sensor provides a sufficiently large bit depth of the images, the behavior of the compression module can be approximated with the distortion characteristic of the compression module when $R_s = 8$ b. If the power scalable video-encoder described in [4] is used, the P-R-D model for the compression module becomes

$$D_c(P_c, R_c, R_s) \approx D_c(P_c, R_c, 8bit)$$
$$= \sigma_c^2 2^{-\lambda_c (P_c/P_{c\max})^{\mu_c} R_c}. \quad (17)$$

In (17), $P_{c\max}$ is the maximum power consumption P_c allocated to the compression module, σ_c^2 is the variance of the encoded video, λ_c is the parameter that characterizes encoder efficiency, and μ_c is the parameter that characterizes dynamic voltage scaling of the encoder, and we set $\mu_c = 2/5$ as in [4].

In this paper, we restrict our attention to higher bit depths of the images and low target rates where the compression distortion dominates the sensor distortion. In these settings, any additional attenuation of the sensor distortion in the compression module

is not critical for the model. Alternative more accurate models that illustrate the effects of compression module can be readily used without significantly affecting the optimization procedure.

In order to obtain power dissipation of the compression module comparable to the power consumption of our imager, we consider a video encoder designed in 0.13 ìm technology. As a reference, we use Amphion's MPEG-4 CS6701 coder that consumes less than 15 mW in 0.18 ìm for compression of color images [23]. If we assume 50% power reduction due to technology scaling and similar reduction due to the fact that we process grayscale images, we set the compression module power to 3.5 mW.

3) Wireless Transmission: To describe the distortion caused by the transmission module, we apply the analysis [5] for video transmission. The analysis considers the distortions caused by the loss of both intra and inter macroblocks. The transmission distortion D_t is related to the packet error probability p and the video characteristics given as

$$D_t = \theta \frac{p}{1-p} \bar{F} d \quad (18)$$

where

$$\theta = \frac{a}{(1-b+b\xi)} \quad (19)$$

and $\bar{F}_d$ is the average value of the frame difference over the whole video scene, ξ is the intra refreshing rate of the video frames, a is a constant that indicates the amount of information discarded by the video coding algorithm, and b is a constant that indicates the motion randomness. The simulation results in [5] verify this distortion model with a prediction error of less than 5%.

For the case of no retransmissions on packet errors, the packet error rate is related to the packet size L and the probability of bit error p_b as

$$p = 1 - (1 - p_b)^L. \quad (20)$$

The probability of bit error p_b is determined by the modulation used along with the transmission power. Several WPAN standards (IEEE 802.15) employ phase-shift-keying (PSK) modulation techniques. In this paper, we assume a Rayleigh flat-fading channel with BPSK modulation, since it is also a good approximation of GFSK (and GMFK) used in IEEE 802.15.1 standard (Bluetooth) up to 1 Mb/s [24]. The probability of bit error is then

$$p_b = \frac{1}{2} \left[1 - \sqrt{\frac{\frac{E_b}{N_0}}{1 + \frac{E_b}{N_0}}} \right] \quad (21)$$

where E_b is the received energy per bit and N_0 is the noise power spectral density. The received energy per bit is determined by the power used for transmission P_t, transmission rate R_t, and the path loss model of the signals. Assuming free-space path loss model, unity antenna gains and unity system loss factor, the received energy per bit can be expressed as

$$E_b = \frac{P_t \lambda^2}{R_t (4\pi d)^2} \quad (22)$$

$\sum\Delta$ Image Sensor		Compression		Wireless Transmission	
P_{smax}	1.35 mW	P_{cmax}	3.5 mW	d	20 m
OSR_{max}	75	μ_c	2/5	λ	1/8 m
σ_s^2	3900 MSE	λ_c	5	L	96 byte
α	1.7×10^{-4}	σ_c^2	150 MSE	N_0	4×10^{-18} W/Hz
β	0			$\overline{F_d}$	200 MSE
γ	0.72			θ	0.8

Fig. 10. Optimum transmission rate at maximum power consumption.

Fig. 9. Minimum distortion of the WVS as a function of the total power consumtion and target bit rate.

Fig. 11. Minimum distortion of the WVS as a function of the total power consumption. Two cases are illustrated, when the image sensor is included and excluded from the power control algorithm.

where λ is the wavelength and d is the distance between transmitter and receiver in meters [25].

In this analysis, we neglect the nonstationary characteristics of the video data and select average video sequence parameters to reflect the typical behavior shown by the analysis in [2] and [4]. The simulation parameters used are summarized in Table I.

B. Simulation Results

Using our P-R-D optimization framework, we analyze the dependence of the minimum system distortion on the available power budget. We apply a pattern search algorithm in MATLAB to find the minimum of the objective function (1) under constraint (2). The optimization procedure is repeated for different target rates. The dependence of the minimum total distortion on the available power budget and target bit rate is shown in Fig. 9. As seen in Fig. 9, the compression distortion dominates at lower target rates; however, at higher target rate the received energy per bit is reduced, which produces higher transmission distortions. Hence, there is an optimal rate that minimizes the target distortion. For maximum power consumption, the optimum rate is found to be close to 96 kbps as shown in Fig. 10. In further analysis, we assume that the target rate is fixed to 96 kbps.

Next, we consider the scenario where the image sensor is not incorporated in the power control and a fixed amount of power is
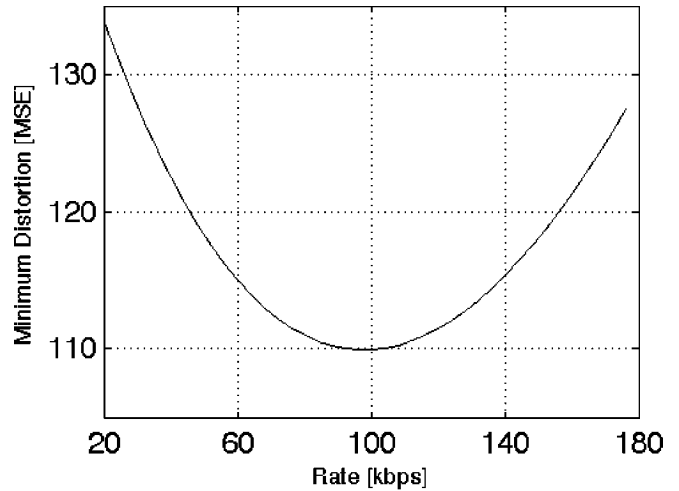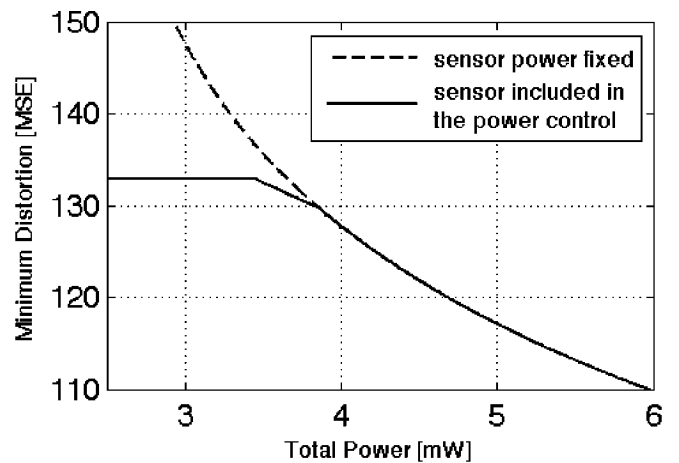
allocated to the image sensor ($P_{s\,max}$). The optimization procedure is applied only to the compression and transmission modules (objective function now has two variables, P_c and P_t). Numerical results from the optimization procedure are shown in Fig. 11. When the image sensor is not incorporated in the power control, the sensor will work at full capacity even at low power levels. When the total power budget decreases below 50% of the maximum value, there is not enough power for compression and transmission operations and, hence, no information about the video content is available (the distortion reaches the maximum value σ_c^2). However, if the image sensor is included in the optimal power allocation procedure (Figs. 9 and 11), the control algorithm limits the distortion at low power values: at $0.5 \times P_{s\,max}$, a WVS can still produce images. On the other hand, when the image sensor is included in the optimization procedure, at low power values, less power can be used to achieve the same level of distortion. Thus, the lifetime of the nodes can be increased.

In Fig. 12, the power allocation among the individual modules as a function of the total power budget is shown for both cases.
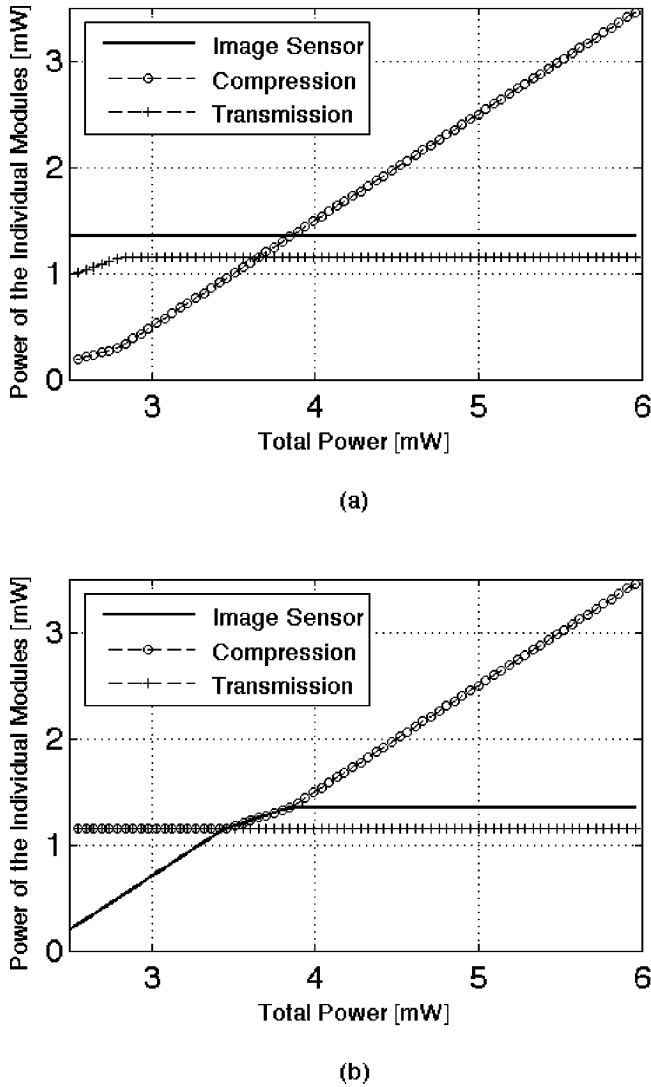
Fig. 12. Distribution of power budget among the sensor, compression, and trasmission modules: (a) sensor power fixed and (b) sensor included in the power control.
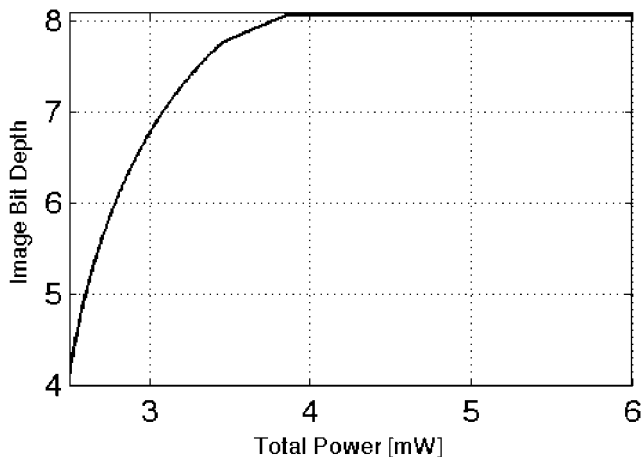

Fig. 13. Adaptive sensing: bit depth of the images captured by the image sensor as a function of the total power budget.

When the image sensor is included in the optimization, by assigning fewer bits to the images as the power budget decreases,

the power from the image sensor is redistributed in the compression module. Fig. 13 illustrates how the bit depth of the captured images reduces as the available power budget decreases. By assigning fewer bits to the images, the image sensor indirectly increases the efficiency of the compression module. The bit depth of the images is traded for improved compression of the spatio-temporal characteristics of the video sequence. In addition, since the ideal $\Sigma\Delta$ modulator doubles the power for each additional 1.5 bit of resolution, a significant influence on the system performance might be expected when the $\Sigma\Delta$ imager is used for high resolution imaging.

V. CONCLUSION

In this paper, we consider optimal power allocation for wireless video sensors (WVSs) containing power and rate scalable image sensor subsystem. We describe an image sensor with a power-rate-distortion (P-R-D) characteristic and use it to build a comprehensive P-R-D framework for WVSs optimization. We derive a P-R-D model for a $\Sigma\Delta$ image sensor and demonstrate how the image bit depth can be adjusted to maximize the overall video quality of the WVS under power constraint. Results from the P-R-D analysis of a WVS indicate that the proposed optimization method can be exploited to prolong the lifetime of the sensor nodes.

In the future work, the design and characterization of a video encoder that exploits large range of the input image bit depth can be investigated. Also, in order to quantify the tradeoff between the image bit depth and other video characteristics under power constraint, a distortion error metric taking into account the response of the human visual system should be examined and compared to the mean square error distortion metric.

REFERENCES

[1] Z. He and S. K. Mitra, "From rate-distortion analysis to resource-distortion analysis," *IEEE Circuits Syst. Mag.*, vol. 5, no. 3, pp. 6–18, Sep. 2005.
[2] Z. He and D. Wu, "Resource allocation and performance analysis of wireless video sensors," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 16, no. 5, pp. 590–599, May 2006.
[3] Z. He, Y. Liang, L. Chen, I. Ahmad, and D. Wu, "Power-rate-distortion analysis for wireless video communication under energy constraints," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 15, no. 5, pp. 645–658, May 2005.
[4] Z. He, W. Cheng, and X. Chen, "Energy minimization of portable video communication devices based on power-rate-distortion optimization," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 18, no. 5, pp. 596–608, May 2008.
[5] 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, no. 6, pp. 511–523, Jun. 2002.
[6] M. Marijan, W. Heinzelman, G. Sharma, and Z. Ignjatovic, "Optimal resource allocation for wireless video sensors with power-rate-distortion model of imager," in *Proc. IEEE Int. Midwest Symp. Circuits and Systems*, Aug. 2009, pp. 256–259.
[7] P. Pirsch, M. Berekovic, H.-J. Stolberg, and J. Jachalsky, "VLSI architectures for MPEG-4," in *Proc. Int. Symp. VLSI Technology, Systems, and Applications*, Oct. 2003, pp. 208–212.
[8] O. Silven, T. Rintaluoma, and K. Jyrkka, "Implementing energy efficient embedded multimedia," in *Proc. SPIE*, 2006, vol. 6074, Multimedia on Mobile Devices II, pp. 61–70.
[9] A. J. P. Theuwissen, "CMOS image sensors: State-of-the-art," *Solid State Electron.*, vol. 52, pp. 1401–1406, 2008.
[10] Digital Pixel System Technology Pixim Inc., 2008 [Online]. Available: http://www.pixim.com/assets/files/product_and_tech/Digital_Pixel_System_Technology_White_Paper_June_2_08.pdf

[11] L. G. McIlrath, "A low-power low-noise ultrawide-dynamic-range CMOS imager with pixel-parallel A/D conversion," *IEEE J. Solid-State Circuits*, vol. 36, no. 5, pp. 846–853, May 2001.

[12] A. Mahmoodi and D. Joseph, "Pixel-level delta-sigma ADC with optimized area and power for vertically-integrated image sensors," in *Proc. IEEE Int. Midwest Symp. Circuits Syst.*, Aug. 2008, pp. 41–44.

[13] D. Maricic, Z. Ignjatovic, and M. F. Bocko, *Low-Power, High-Dynamic Range CMOS Image Sensor Employing Pixel-Level Oversampling Sigma-Delta Analog-to-Digital Conversion*, unpublished.

[14] M. Bigas, E. Cabruja, J. Forest, and J. Salvi, "Review of CMOS image sensors," *Microelectron. J.*, vol. 37, pp. 433–451, 2006.

[15] A. E. Gamal and H. Eltoukhy, "CMOS image sensors," *IEEE Circuits Devices Mag.*, vol. 21, no. 3, pp. 6–20, May-Jun. 2005.

[16] H. Tian, B. Fowler, and A. E. Gamal, "Analysis of temporal noise in CMOS photodiode active pixel sensor," *IEEE J. Solid-State Circuits*, vol. 36, no. 1, pp. 92–101, Jan. 2001.

[17] P. M. Aziz, H. V. Sorensen, and J. vn der Spiegel, "An overview of sigma-delta converters," *IEEE Signal Process. Mag.*, vol. 13, no. 1, pp. 61–84, Jan. 1996.

[18] Noise Sources in CMOS Image Sensors Hewlett-Packard Company, 1998 [Online]. Available: http://www.stw.tu-ilmenau.de~ff/beruf_cc/cmos/cmos_noise.pdf

[19] R. I. Hornsey, Part III: Noise in Image Sensors University of Waterloo [Online]. Available: http://www.cse.yorku.ca/visor/pdf/CMOS_3.pdf

[20] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Hoboken, NJ: Wiley, 2006.

[21] B. Widrow, I. Kollar, and M.-C. Liu, "Statistical theory of quantization," *IEEE Trans. Instrum. Meas.*, vol. 45, no. 2, pp. 353–361, Apr. 1996.

[22] S. Borkar, "Design challenges of technology scaling," *IEEE Micro*, vol. 19, no. 4, pp. 23–27, Jul.–Aug. 1999.

[23] Amphion Accelerator Cores Deliver Low-Power MPEG-4 Encoder Solutions for Digital Video Sillicon Design Design & Reuse, 2002 [Online]. Available: http://www.design-reuse.com/news/4260/amphion-accelerator-cores-power-mpeg-4-encoder-solutions-digial-video-silicon-design.html

[24] S. Galli, D. Famolari, and T. Kodama, "Bluetooth: Channel coding considerations," in *Proc. IEEE Vehicular Technology Conference*, Milan, Italy, May 17–19, 2004, vol. 5, pp. 2605–2609.

[25] T. S. Rappaport, *Wireless Communications: Principles and Practice*. Upper Saddle River, NJ: Prentice-Hall, 1996.

Mališa Marijan (S'09) received the B.S. degree in electrical engineering from the University of Belgrade, Serbia, in 2005, the M.S. degree in electrical engineering from the University of Rochester, Rochester, NY, in 2008, and is currently working toward the Ph.D. degree in electrical engineering at the University of Rochester.

His research interests include signal processing, application of information theory and communications to mixed-signal circuits design and modeling, A/D conversion methods, and image sensor architectures.

Ilker Demirkol (S'01–M'02) received the B.S. (with honors), M.S., and Ph.D. degrees in computer engineering from Bogazici University, Istanbul, Turkey, in 1998, 2002, and 2008, respectively.

Currently, he is a Postdoctoral Researcher at the University of Rochester, NY. He worked as a database, system, and network engineer from 1997 to 2004. He was a research and teaching assistant at Bogazici University, Computer Engineering Department, from 2004 to 2008. His research interests include the areas of wireless communications, wireless ad hoc and sensor networks, and the optimization of communication networks.

Danijel Maričić (S'08–M'10) received the B.S. degree in electrical engineering from the University of Belgrade, Serbia, in 1998, the M.S. degree in electrical engineering from the University of Rochester, Rochester, NY, in 2007, and is currently working toward the Ph.D. degree in electrical engineering at the University of Rochester.

His research interests include low power and high dynamic range CMOS image sensors and power efficient image processing architectures.

Gaurav Sharma (SM'00) received the B.E. degree in electronics and communication engineering from the Indian Institute of Technology Roorkee (formerly University of Roorkee), India, in 1990; the M.E. degree in electrical communication engineering from the Indian Institute of Science, Bangalore, India, in 1992; and the M.S. degree in applied mathematics and the Ph.D. degree in electrical and computer engineering from North Carolina State University (NCSU), Raleigh, in 1995 and 1996, respectively.

From August 1992 through August 1996, he was a research assistant at the Center for Advanced Computing and Communications, ECE Department, NCSU. From August 1996 through August 2003, he was with Xerox Research and Technology, Webster, NY, initially as a member of research staff and subsequently at the position of principal scientist. Since Fall 2003, he has been an Associate Professor at the University of Rochester in the Department of Electrical and Computer Engineering and in the Department of Biostatistics and Computational Biology. He is also the Director for the Center for Electronic Imaging Systems (CEIS), a New York state funded center for promoting joint university-industry research and technology development, which is housed at the University of Rochester. His research interests include multimedia security and watermarking, color science and imaging, genomic signal processing, and image processing for visual sensor networks. He is the editor of the *Color Imaging Handbook* (CRC, 2003).

Dr. Sharma is a member of Sigma Xi, Phi Kappa Phi, Pi Mu Epsilon, IS&T, and the signal processing and communications societies of the IEEE. He was the 2007 chair for the Rochester section of the IEEE and served as the 2003 chair for the Rochester chapter of the IEEE Signal Processing Society. He currently serves as the vice chair for the IEEE Signal Processing Society's Image and multidimensional signal processing (IMDSP) technical committee and is an advisory member of the IEEE Standing committee on Industry DSP. He currently serves as an associate editor the *Journal of Electronic Imaging* and in the past has served as an associate editor for the IEEE TRANSACTIONS ON IMAGE PROCESSING and the IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY.

Željko Ignjatović (S'03–M'04) received the B.S. degree in electrical engineering from University of Novi Sad, Serbia, and the M.S. and Ph.D. degrees in electrical engineering from University of Rochester, NY, in 1999, 2001, and 2004, respectively.

He is currently an Associate Professor of Electrical and Computer Engineering at the University of Rochester. His areas of interest include analog circuit design and development of A/D converters, image sensor architectures, radar and ultrasound imaging techniques, and related signal processing methods.