Optimal Resource Allocation for Wireless Video Sensors with Power-Rate-Distortion Model of Imager

Malisa Marijan, Wendi Heinzelman, Gaurav Sharma, and Zeljko Ignjatovic

Department of Electrical and Computer Engineering,
University of Rochester,
Rochester, NY 14627 USA
Email: {marijan; wheinzel; gsharma; ignjatov}@ece.rochester.edu

Abstract—**We consider optimal resource allocation for wireless video sensors (WVSs), including the image sensor subsystem into the system analysis. By assigning a power-rate-distortion (P-R-D) characteristic for the image sensor, we build a comprehensive P-R-D framework for WVS optimization. Within the scope of the developed framework, we solve the problem of how to allocate power among the image sensor, compression, and transmission modules of a WVS, to achieve the optimal reconstructed video quality under power and rate constraints. To demonstrate the optimization method, we further establish a P-R-D model for an image sensor based on pixel level sigma-delta ($\sum\Delta$) image sensor design. A P-R-D performance analysis for a WVS verifies that including the image sensor in the system optimization procedure can improve the overall video quality and prolong the life-time of the wireless video network.**

I. INTRODUCTION

In a typical video communication system, one of the major problems in system optimization is control of the system performance under bandwidth constraints. Rate-distortion (R-D) theory is usually applied to analyze system behavior within bandwidth limitations. 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 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 sensor networks operate under limited power supply. The available power plays an important role in the resulting video quality as well as in the life-time of the system. The problem of power allocation among the image sensor, compression, and transmission modules in the sensor nodes is an important problem in the system optimization. Because 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 R-D analysis for video sensor networks has been addressed in [1]. It is suggested that for optimal resource allocation, classical R-D analysis has to 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 of power consumption and R-D analysis must be merged together [1], [2].

In wireless local area networks (50-100m 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]. An optimization framework based on the P-R-D model of the video encoder is then used in [2] to analyze the power trade-off between the video encoding and wireless data transmission modules. However, simulation results from [4] indicate that in a 10-20m communication range, the camera can consume 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 trade-off of an image sensor depends strongly on the specific sensor design, which makes attempts to derive general specifications rather difficult.

In a personal area network where communication is typically performed within 10m range, the power requirements for the transmission module are further reduced. In this case, resource allocation cannot be performed in an optimal way without incorporating the image sensor in the system analysis. We believe that if the compression module is designed to accept different data rates from the image sensor (pixel bit depth), controlling the output rate of the image sensor will

lead to improved power allocation. For example, if the sensor is included in power control, the power consumption and rate of the sensor can be reduced as the power budget decreases. Reducing the power consumption of the image sensor can increase overall video quality because more power can be allocated to the compression and transmission modules. In addition, a lower rate of the image sensor means that distortions of the compression module can be decreased due to the need for lower compression ratios.

In order to build a comprehensive P-R-D framework for WVSs, we propose a P-R-D characteristic for an image sensor. This approach allows establishing an optimal solution for the power allocation problem for WVS containing arbitrary image sensor with changeable output rates. In Section II, we define and solve the optimization problem for WVS. Furthermore, an analytic P-R-D model for a sigma-delta ($\sum\Delta$) imager is proposed in Section III. In Section IV, we perform P-R-D analysis for a WVS and discuss the optimization results. Our work is summarized in Section V.

II. OPTIMAL POWER ALLOCATION

In this section, the problem of optimal power allocation among the image sensor, compression, and transmission modules in a WVS is analyzed. Although we consider an image sensing application, the derived optimization method can also be utilized for other types of sensors.

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 will 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 defined as: *How should power P be allocated among the image sensor, compression, and transmission modules to minimize the overall distortion introduced for a given rate R?*

The most common measure for video quality is end-to-end distortion D, which represents the mean square error between the original and received picture frames [2]. Distortion in each module can be defined as the mean squared difference between the output and input frame of the module. Distortions introduced by the three modules are assumed to be independent, conditioned on given data rate R_s of the image sensor, so overall distortion is the sum of the distortions over all three subsystems.

We assume that no channel coding is applied so that the bit rates R_c and R_t are equal to the target bit rate R (Fig.1). We also assume that 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 $\sum\Delta$ imager discussed in Section III), so that the image sensor distortion D_s depends only on the power P_s. 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 output rate of the image sensor determines the required compression ratio (since
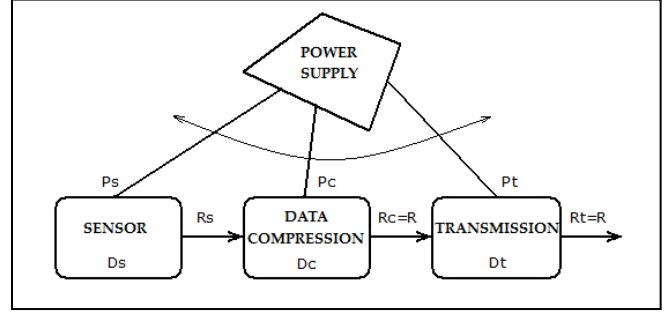

Figure 1.   Wireless video sensor: power has to be optimaly allocated among the image sensor, compression, and transmission modules.

the output rate of the compression module must be the target bit rate R). Therefore, compression distortion is also a function of the image sensor rate, i.e., the image sensor power consumption. The optimization problem can be mathematically formulated as:

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

subject to

$$P_s + P_c + P_t = P. \quad (2)$$

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 is shown in [2].

The optimal solution can be calculated using the Lagrange multiplier method. In this method, the optimization problem becomes the problem of finding the minimum of the unconstrained function $\Lambda = \Lambda(P_s, P_c, P_t, \lambda)$ that is given in the form:

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

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:

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

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

Equalities (4) imply that for an optimal solution P_s^*, P_c^*, P_t^*, the tangents on the distortion functions $D_s(P_s) + D_c(P_s, P_c^*)$, $D_c(P_s^*, P_c)$, and $D_t(P_t)$ must have the same slope.

III. P-R-D Model for ∑Δ Image Sensor

In this section, a P-R-D model for a specific image sensor is developed. A digital pixel sensor (DPS) architecture is considered, where analog-to-digital conversion is performed simultaneously at every pixel location. Due to a high degree of parallelism, the requirement for speed of conversion is relaxed, which translates into low power consumption. Low power consumption makes DPS suitable for applications that are limited in power supply, such as WVSs.

A ∑Δ imager is a DPS architecture performing ∑Δ data conversion at each pixel site. During the exposure time of the imager, a ∑Δ modulator reads charge from the photodiode OSR (oversampling ratio) number of times, and each value is characterized with one bit (see Fig. 2). A decimation filter is then applied in order to convert the oversampled single-bit stream to the multi-bit samples at the frame rate.

A ∑Δ imager is suitable for establishing a P-R-D model and it allows relatively simple power control. Power, rate, and performance of the imager can be controlled by changing the number of pixel values that are read every second. In other words, OSR can be used as a control parameter for P-R-D modeling and resource allocation, while keeping other design parameters such as power supply voltage fixed.

To characterize the distortion within each pixel of a ∑Δ image sensor, we derive a R-D model for a pixel with respect to the quantization operation of the ∑Δ converter. A R-D characteristic of quantization is used to describe the ∑Δ conversion, where the rate corresponds to the effective number of bits ($ENOB$) provided by the ∑Δ converter. We also assume that signals read from the photodiodes have a Gaussian distribution with variance σ_{ij}^2, for each pixel at position defined by indices i and j. The overall distortion introduced by the ∑Δ imager is then calculated as the average distortion introduced by the pixels of the frame:

$$D_s = \frac{1}{NM} \sum_i^N \sum_j^M \sigma_{ij}^2 2^{-2ENOB} = \sigma^2 2^{-2ENOB}, \quad (6)$$

where NxM is the number of pixels in the image sensor array, and σ^2 represents the variance of the picture frame. In addition, the maximum distortion cannot be greater than σ^2. The actual output bit rate of the sensor in bits per pixel is an integer value, and hence the output rate of the sensor is $ENOB$ rounded to the next largest integer value, i.e., $R_s = \lceil ENOB \rceil$.

In order to simplify the calculation of the $ENOB$ of the ∑Δ imager, we consider a linear model of the ∑Δ modulator and that the decimation filter is an ideal low pass filter. In addition, we assume that the quantization noise is the only source of signal distortion. The $ENOB$ at the output of an ideal ∑Δ converter can be readily calculated as shown in (7).

$$ENOB = \frac{1}{2} \log_2 \left(\frac{3OSR^3}{\pi^2} \right). \quad (7)$$

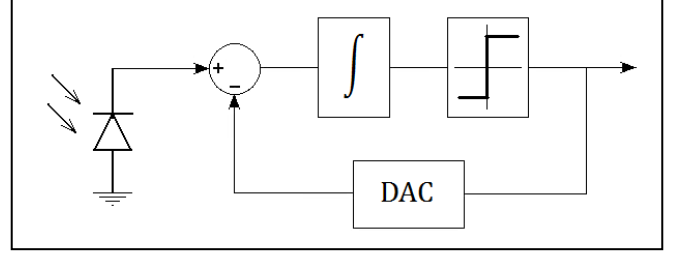Figure 2.   A ∑Δ imager: charge from each photodiode is read by a ∑Δ modulator consisting of an integrator, a single bit comparator and a digital-to-analog converter.

Power consumption of the imager contains two components: static and dynamic. Static power consumption of a fully-digital pixel sensor often can be neglected with respect to the dynamic power. Dynamic power consumption is, to the first approximation, observed to be a linear function of the sampling frequency. Thus, the following can be written:

$$P_s = \frac{OSR}{OSR_{max}} P_{smax}, \quad (8)$$

where P_{smax} is the image sensor power consumption at the maximum output rate R_{smax}, i.e., at the maximum oversampling ratio OSR_{max}. Combining (6)-(8), we obtain a P-D characteristic for the ∑Δ image sensor:

$$D_s = \frac{\pi^2 \sigma^2}{3OSR_{max}^3} \left(\frac{P_{smax}}{P_s} \right)^3. \quad (9)$$

IV. P-R-D Performance Analysis

In order to verify the proposed P-R-D framework, we consider an example of a WVS network. We assume that WVSs are deployed in a small area (e.g. in a five meter range) and supplied with ∑Δ imagers designed to capture 8-bit grayscale images in CIF (352 x 288) format at 15 f/s. The video sequence is than compressed to reduce the bandwidth requirements, and the compressed bit stream is transmitted over the wireless medium at a rate of 1Mb/s.

To describe the behavior of the compression module, we apply the P-R-D characteristic for power-scalable video encoding proposed in [2]. However, to account for the fact that the compression ratio depends on the output rate of the image sensor, we introduce a scaling factor R_{smax}/R_s in the model, as described by the following expression:

$$D_c = \sigma^2 2^{-\lambda_c \frac{R_{smax}}{R_s} R \left(\frac{P_c}{P_{cmax}} \right)^{\mu_c}}. \quad (10)$$

In (10), R is the data rate in bits per pixel, P_{cmax} is the maximum power consumption P_c allocated to the compression module, and λ_c and μ_c are model parameters. To simplify the analysis, we assume that the picture frame variance is the

same for all three modules and equals to σ^2. We further neglect the non-stationarity of the video data and select model parameters to resemble typical behavior shown by the analysis in [2]. In particular, we set $\lambda_c = 3$ and $\mu_c = 2/3$.

We assume that video content is not packetized and that transmission distortions come from bit errors that occur with probability p_b. We propose a distortion model for the transmission module in the form $D_t = \sigma^2 p_b / (1 - p_b)$, so that for $p_b = 0.5$ transmission distortion has a maximum value σ^2, and for $p_b = 0$ there is no distortion. To achieve a rate of 1Mb/s, we assume that binary DPSK modulation is applied. Consequently, the bit error probability is $p_b = \frac{1}{2} e^{-E_{br}/N_0}$, where E_{br} is the received energy per bit and N_0 is the noise power spectral density. The distortion model can be rewritten in the form:

$$D_t = \frac{\sigma^2}{2 e^{\lambda_t \frac{P_t}{R}} - 1}, \qquad (11)$$

where the system parameter λ_t relates the received energy per bit to the transmitted energy P_t / R. We assume that the system is designed to provide bit error probability on the order of 10^{-5}.

The P-D characteristics for the image sensor, compression, and transmission modules of the WVS are shown in Fig. 3. The characteristics of the compression module are provided for two sensor output values, illustrating how distortions of the compression module can be reduced by decreasing the output rate of the image sensor. The shape of the P-D characteristic for the $\sum \Delta$ imager implies that the sensor should not be run at full capacity. The output rate of the $\sum \Delta$ image sensor should be lowered to reduce compression and transmission distortions, and then slightly changed if the available power budget decreases.

Using our P-R-D optimization framework, we analyze how minimum system distortion depends on the total power
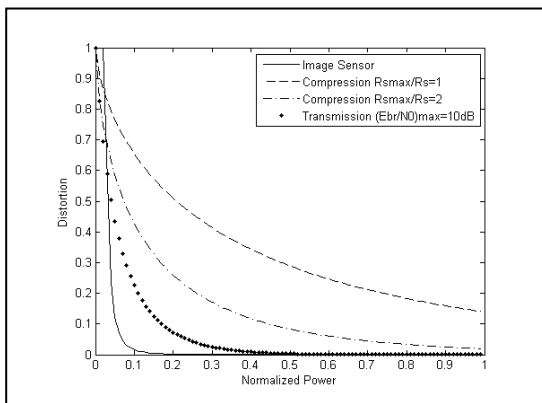
Figure 3. Power-distortion characteristics for sensing, compression, and transmission. Distortion is normalized to the maximum value σ^2, and power consumption of each module is normalized to the maximum power allocated to the corresponding module (without optimization).
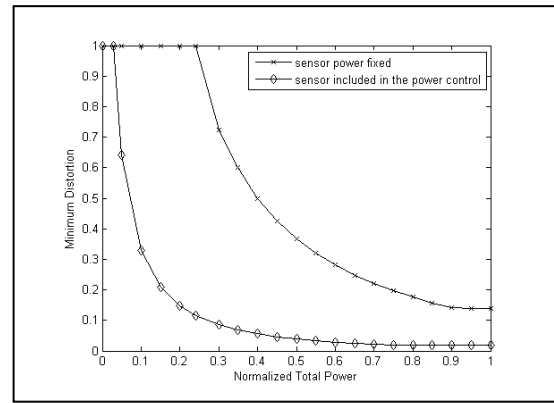
Figure 4. Minimum distortion of the WVS as a function of total power consumption.

consumption. First, we consider the case when the fixed amount of power is allocated to the image sensor, and the optimization procedure is applied only to the compression and transmission modules. Numerical results are shown in Fig. 4 for the case where the maximum total power is allocated to the modules in a ratio of $P_s : P_c : P_t = 1:3:1$. When the image sensor is not incorporated in the power control, it will work at full capacity even at low power levels. At 20% of the maximum total power, there is not enough power for compression and transmission operations, and hence no information about the video content is available. However, if the image sensor is included in the optimal power allocation (Fig. 4), at 20% of the maximum total power, a WVS can still produce images with relatively low distortion.

V. CONCLUSION

In this paper, we introduce the concept of describing an image sensor with a P-R-D characteristic, and we use this to build a comprehensive P-R-D framework for optimal resource allocation and power control for WVSs. Results from the P-R-D analysis of a WVS show that the proposed optimization method can be exploited to minimize system distortions and to prolong the life-time of a wireless video network.

REFERENCES

[1] Z.He and S.K. Mitra, "From rate-distortion analysis to resource-distortion analysis," IEEE Circuits and System Magazine, vol. 5, pp. 6-18, September 2005.

[2] Z. He and D. Wu, "Resource allocation and performance analysis of wireless video sensors," IEEE Transactions on Circuits and Systems for Video Technology, 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 Transactions on Circuits and Systems for Video Technology, 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 Transactions on Circuits and Systems for Video Technology, vol. 18, No.5, pp. 596-608, May 2008.

[5] P.M. Aziz, H.V. Sorensen, and J. vn der Spiegel, "An overview of sigma-delta converters," IEEE Signal Processing Magazine, vol. 13, No. 1, pp. 61-84, Januar 1996.

[6] Z.Ignjatovic and M.Bocko, "A 0.88nW/pixel, 99.6dB linear-dynamic-range fully-digital image sensor employing a pixel-level sigma-delta ADC," VLSI Simposium, 2006.