JOINT ENERGY OPTIMIZATION OF VIDEO ENCODING AND TRANSMISSION

Ziyu Ye*, Rana Hegazy*, Wei Zhou†, Pamela Cosman*, Larry Milstein*
*ECE Department, University of California, San Diego, La Jolla, CA92093, United States
{ziy076,rhegazy,pcosman,milstein}@eng.ucsd.edu
†School of Elctronics and Information, Northwestern Polytechnical University, Xian, 710072, China
zhouwei@nwpu.edu.cn

***Abstract*—Disposable wireless video sensors have many potential applications but are subject to stringent energy constraints. We studied the minimization of end-to-end distortion under an total energy constraint, by means of optimizing FEC code rate, number of source bits, and energy allocation between video encoding and wireless transmission. A two-step approach is employed. First, the FEC rate is optimized by exhaustive search. Then a binary-search-based algorithm is proposed to optimize the energy allocation and number of source bits. Experiments show that the algorithm achieves a PSNR gain up to 1dB over some reasonable baselines. A simpler suboptimal algorithm is also tested and exhibits similar performance.**

I. Introduction

Video surveillance in remote areas can be useful for monitoring pipelines or reservoirs against terrorists, border patrol, measuring snowpack levels or other environmental conditions, or studying endangered species. While remote areas may lack infrastructure, many cheap disposable wireless video sensors can be deployed to form a flexible surveillance system. Whether using a fixed battery or energy harvesting, such sensors need to be energy efficient. Most of the energy consumption comes from three operations: video capture, encoding, and transmission. This paper focuses on the joint optimization of video encoding and transmission.

Many papers have considered joint optimization of video encoding and wireless transmission in the sense of bit allocation, but fewer involve energy constraints. In [1], the minimization of video transmission energy over source rate and transmitting power is studied. The source rate and transmitting power are jointly constrained by a fixed distortion budget. In [2], minimization of the total energy consumed by both video encoding and transmission is studied. A rate-complexity-distortion (RCD) model is used to introduce the video encoding energy (source energy) into discussion. Each source bit is assumed to be allocated a sufficient amount of transmission energy so that channel errors are negligible. One limitation of this work is that the transmission model is simplified; no details about transmission, including the optimality of the FEC code rate, are considered.

We study the joint optimization of video encoding and transmission for wireless video sensors operating in a remote area. In this scenario, energy is limited but bandwidth is abundant. Under an energy constraint, we consider the minimization of the end-to-end distortion, by jointly optimizing energy allocation, source rate, and FEC code rate. We study the optimization problem using a specific transmission model, and an empirical RCD model describing the behavior of the HEVC/H.265 video encoder. This paper is arranged as follows. Section II formulates the optimization problem. Sections III and IV develop the RCD and transmission models. Section V introduces two algorithms that efficiently solve the optimization problem, and Section VI presents simulation results.

II. Problem Formulation

Consider a scenario where wireless video sensors monitor a a remote area. A drone flies over periodically to collect data. When the drone is close (e.g., 100m), the sensor transmits the data. In daily operation, the sensors capture/encode video, add FEC, divide the bitstream into packets, and store the data in a buffer (Fig. 1). At the receiver, after the FEC code is decoded, packets containing uncorrectable errors are discarded, and error concealment is employed in the video decoder.

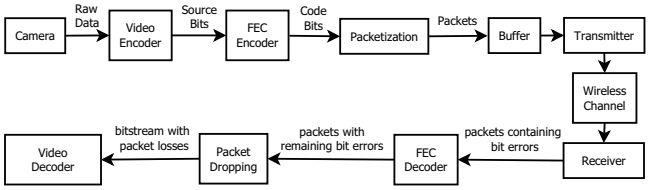

Fig. 1. System diagram.

Video quality is measured by the end-to-end distortion, which is the mean square error (MSE) between the original video and the decoder reconstruction. Several papers (e.g., [3]) showed that this can be approximated by the sum of the source distortion, D_s, (MSE between the original video and the encoder reconstruction) and channel distortion, D_{chan} (MSE between the encoder and decoder reconstructions).

We constrain the total energy spent on a unit length (1 second) of video. This total energy is the sum of energies consumed by encoding and transmission: $E_{TX} = E_s + E_t$. Energy for FEC encoding and digital modulation are small compared to that of the analog synthesizer & filter [4], and are ignored. Then E_t consists of only the energy consumption of the analog circuits and the energy transmitted. No bandwidth constraint is imposed as the sensors operate in remote areas,

where radio spectrum is sparsely occupied. Transmission time is also not constrained as video delivery is not real time.

The energy constraint gives rise to an allocation problem: If more energy is allocated to encoding, the encoder can find more efficient representations of the video and reduce the D_s. With more energy for transmission, either higher power or more FEC redundancy can be used to reduce the D_{chan}. Another tradeoff concerns the number of source bits (bits generated by the video encoder), denoted by B_s. For any given energy allocation (given E_s, E_t), increasing B_s means D_s is lower, but the increase in number of source bits needs to be compensated by either reducing the number of FEC parity bits or spending less transmission energy on each code bit, making the data more vulnerable to channel noise. These two tradeoffs lead to the following optimization problem, in which we look for the energy allocation and the number of source bits that minimize the end-to-end distortion:

$$\min_{E_s,E_t,B_s} [D_s(Es,B_s)+D_{chan}(E_t,B_s)] \quad s.t. E_s+E_t=E_{TX} \tag{1}$$

Here we consider no explicit constraint on B_s. We assume the buffer has enough space for the compressed video data. Substituting the constraint into the objective function gives

$$\min_{E_s,B_s} [D_s(E_s,B_s)+D_{chan}(E_{TX}-E_s,B_s)] \tag{2}$$

In the following, we will construct the functions $D_s(E_s,B_s)$ and $D_{chan}(E_t,B_s)$, and introduce a procedure to solve (2). We use MFSK modulation for its energy efficiency, Reed-Solomon coding as the FEC, and the HM-16.1.1 HEVC/H.265 standard software (random access main profile, with periodic I-frame refresh and bidirectional B frames).

III. Source Model

In this section, we construct the function $D_s(E_s,B_s)$ in (1). The encoder can vary the quantization parameter (QP) and 5 other configuration parameters (Table I). There are many QP-configuration pairs. With different QPs and configurations, the encoder generates from a given video different numbers of source bits B_s, consuming different amounts of encoding energy E_s, and resulting in different source distortions D_s. Although E_s, B_s, D_s take discontinuous values corresponding to the discrete QPs and configurations, we treat E_s and B_s as continuous variables and D_s as a function of the two, forming a surface over the E_s-B_s plane. By fitting the surface with a parametric model, we model D_s as a function of B_s and E_s.

In a rate-complexity-distortion (RCD) model (e.g., [2], [5]), the authors incorporate energy (or power) considerations into a rate-distortion model, taking the source distortion to be a function of the source rate and source complexity. The source complexity (denoted C_s) is a normalized measurement of the computational complexity of the video encoding process, and is defined as being proportional to the number of CPU clock cycles required to encode a fixed length of video. The complexity is normalized so that it varies within the range

TABLE I
Video encoder parameters.

Name	Meaning	Range used
QP	Quantization parameter	$\{22,\ldots,37\}$
CUD	Maximum coding unit depth	$\{1,2,3,4\}$
TUMD	Maximum transform unit depth	$\{2,3\}$
HME	Hadamard transform	$\{0,1\}$
SR	Search range	$\{8,16,32,64\}$
BSR	Search range for bi-prediction refinement	$\{2,4\}$

0 to 1, with 1 being the complexity of the configuration that requires the heaviest computation (the most CPU clock cycles). According to [2], the complexity C_s can be mapped to the source energy E_s using the model $E_s=\beta\cdot C_s^{\gamma}$, where γ is a constant, and β is the energy consumed by encoding one second of video with the most complex configuration. We use $\gamma=2.5$ as in [2]. With this mapping between C_s and E_s, the RCD model is equivalent to a rate-energy-distortion model, which is the $D_s(E_s,B_s)$ function we need. We build the RCD model by using a parametric model to fit rate-complexity-distortion data obtained by encoding sample videos:

$$D_s=e^{-a_1\cdot B_s+b_1}\cdot(1+e^{-a_2\cdot C_s+b_2})+D_{floor} \tag{3}$$

In this model, exponential functions are ussd to model the decreasing trend of D_s, as in [2]. Other model features are based on observations made on experimental data. Fig. 2 illustrates data obtained by encoding a video with 164 selections of QP and configuration. Each (B_s,C_s,D_s)-point corresponds to a specific pair of QP and configuration. Points with the same QP are connected with a dashed line.

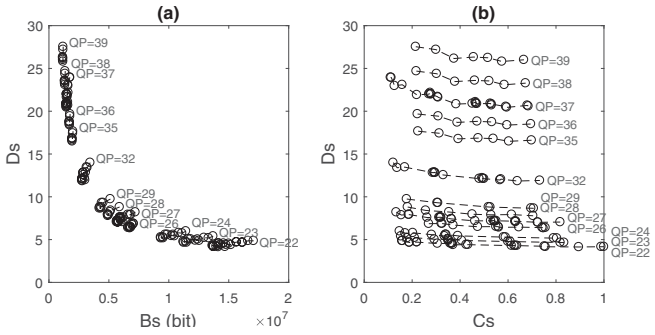

Fig. 2. RCD data for sample video *Traffic*. Points generated with the same QP are connected by a dashed line.

In Fig. 2(a), each little cluster of points labeled by a QP value corresponds to a set of different C_s values. One can observe that when B_s is large, D_s drops to the same floor level despite different values of C_s, which is reflected by D_{floor} in (3). In Fig. 2(b), with different QP used, D_s drops to different levels when C_s becomes large. Since B_s depends almost solely on QP, these facts together suggest that D_s drops to a B_s-dependent floor when C_s becomes large. This B_s-dependent floor is modeled as $e^{-a_1\cdot B_s+b_1}+D_{floor}$ in (3).

IV. Transmission Model

In this section, we derive $D_{chan}(E_t, B_s)$ in (1) in two steps. First, channel distortion D_{chan} is derived as a function of transmission parameters. Second, we constrain both E_t, B_s, and minimize D_{chan} over the transmission parameters.

A. Channel Distortion and Packet Loss Rate

Most works (e.g., [6]) on end-to-end distortion estimation aim at providing information for error-resilient mode decisions, proposing algorithms that use, for example, pixel values and motion vectors to trace the expected distortion on the block level or even the pixel level. Though such algorithms accurately reflect the effect of error concealment and error propagation, they are not suitable for our purposes. First, tracing error propagation requires the motion vector, which depends on the encoding configuration, which is yet to be selected. Second, tracing error propagation on the block, or pixel level requires a considerable amount of computation, conflicting with our goal of energy efficiency.

We approximate the channel distortion as being proportional to the packet loss rate p: $D_{chan} = c \cdot p$, where c is a curve fitting parameter. This assumption is reasonable as sparse packet losses tend to influence disjoint sets of frames and the distortions are additive. In this approximation, we ignore the direct dependence of D_{chan} on the video encoding configuration. Though changing encoding configuration can affect motion estimation and influences the error propagation, we observe that this influence is small when compared to D_{chan}'s dependence on p. Fig. 3 shows the approximation is good despite different configurations used.

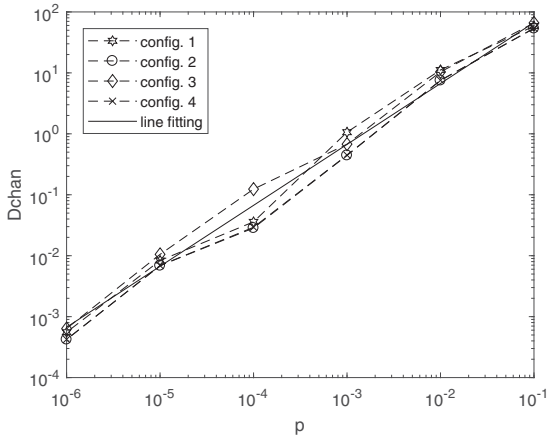

Fig. 3. Channel distortion versus packet loss rate.

We assume the channel coherence time is smaller than the transmission duration of a packet, so packet losses are independent, and the packet loss rate, p, can be written as

$$\begin{aligned} p &= \underset{i}{E}[p_i] = \underset{i}{E}[1-(1-WER)^{N_i}] \\ &\approx \underset{i}{E}[N_i \cdot WER] = \underset{i}{E}[N_i] \cdot WER = N \cdot WER \end{aligned} \tag{4}$$

where p_i is the probability that the i-th packet is lost, N_i is the number of codewords in the i-th packet, N is the average number of codewords in a packet, and WER is the probability that a RS codeword is uncorrectable. It can be shown that $N = (B_s/N_f)/(n \cdot r \cdot m)$, where N_f is the number of frames in 1 second of video, and r is the RS code rate, n is the number of RS code symbols in a codeword, and m is the number of bits represented by a code symbol. WER is:

$$WER = \sum_{j=t+1}^{n} \binom{n}{j} \cdot SER^j \cdot (1-SER)^{n-j}, \tag{5}$$

where $t = \frac{n}{2} \cdot (1-r)$ is the error correction ability, and $SER = 1-(1-P_e)^{\frac{m}{log_2 M}}$ is the probability that a RS code symbol contains error before FEC decoding. Under the assumption of Rayleigh fading and independent MFSK symbol errors, the probability of error in detection of an MFSK symbol is

$$\begin{aligned} P_e = & \sum_{j=1}^{M-1} (-1)^{j+1} \cdot \binom{M-1}{j} \\ & \cdot \frac{1}{(2\sigma^2 \cdot \frac{E_b}{N_0} \cdot log_2 M + 1) \cdot j + 1} \end{aligned} \tag{6}$$

where E_b is transmitted energy per code bit, N_0 is noise power spectral density, σ^2 is half of the second moment of the channel coefficient's amplitude, and M is constellation size.

B. Optimization of FEC code rate

Summarizing the results in IV.A, D_{chan} can be written as

$$D_{chan} = c \cdot \frac{1}{N_f \cdot n \cdot m} \cdot B_s \cdot \frac{1}{r} \cdot WER(E_b, r, M, n) \tag{7}$$

We fix $n = 255$ and $M = 16$ With M, n, m fixed, the term $c \cdot \frac{1}{N_f \cdot n \cdot m}$ is a constant, denoted $const$. For arbitrary E_t and B_s, E_b and r can be optimized as:

$$\min_{E_b, r} \frac{1}{r} \cdot WER(E_b, r), \quad s.t. \ (\frac{E_b}{\eta} + E_{b\_c})/r = \frac{E_t}{B_s} \tag{8}$$

In the constraint, η is the amplifier's efficiency, and E_{b_c} denotes the energy consumed by other analog circuit components for each code bit. Then $E_b/\eta + E_{b\_c}$ is the total transmission energy for each code bit. Since each source bit is encoded into $1/r$ code bits by the FEC encoder, the transmission energy for each source bit is $(E_b/\eta + E_{b\_c})/r$. The constraint reflects the fact that the transmission energy per source bit equals the transmission energy E_t divided by the number of source bits B_s. For notational convenience, we define the transmission energy per source bit as $E_{t\_sb} = E_t/B_s$ and its normalized version as $E^{(n)}_{t\_sb} \triangleq E_{t\_sb} \cdot \frac{\sigma^2}{N_0}$.

For the parameters η and E_{b_c}, following the example of [4], we assume a class-B amplifier is used, for which $\eta = 0.75$. E_{b_c} can be evaluated as $P_c \cdot \frac{M}{B \cdot log_2 M}$, where P_c is the power of the circuit components, B is the total bandwidth and $\frac{M}{B \cdot log_2 M}$ is the time it takes to transmit each code bit. P_c is set to $52.5mW$, which is its value in [4], and we assume $B = 20MHz$. Then $E_{b\_c} = 1.05 \times 10^{-8} J$.

The optimization over E_b and r can be solved numerically by exhaustive search over values of r. The minimized $\frac{1}{r}\cdot WER$ and optimal r are plotted versus $E_{t\_sb}^{(n)}$ in Fig. 4.

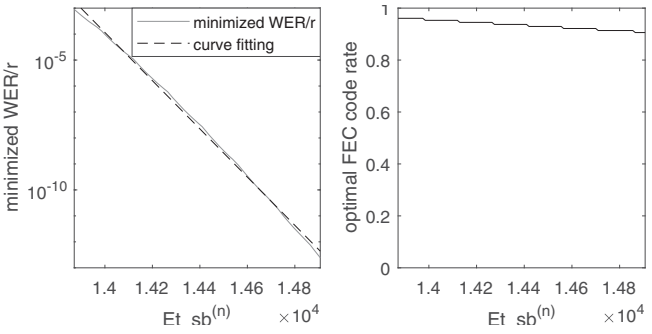

Fig. 4. Minimized WER/r and optimal r as functions of $E_{t\_sb}^{(n)}$.

Fig. 4 shows that the minimized $\frac{1}{r}\cdot WER$ can be approximated as:

$$\frac{1}{r}\cdot WER = e^{-a\cdot E_{t\_sb}^{(n)}+b} \tag{9}$$

where a,b are curve fitting parameters. The parameter values depend on n, M, and E_{b_c}. In different applications, these settings can be different, and the fitted parameters will change. Although the curve is fitted in a narrow range of $E_{t\_sb}^{(n)}$, in the vertical axis, it covers a significant range of $\frac{1}{r}\cdot WER$, which maps to a wide enough range of packet loss rate p. Substituting the approximation of $\frac{1}{r}\cdot WER$ into the expression of D_{chan}, the closed-form function $D_{chan}(E_t, B_s)$ is obtained:

$$D_{chan} = const\cdot B_s\cdot e^{-a\cdot E_{t\_sb}^{(n)}+b} \tag{10}$$

where

$$E_{t\_sb}^{(n)} = \frac{E_{TX}-E_s}{B_s}\cdot\frac{\sigma^2}{N_0} \tag{11}$$

V. Algorithms

Binary-search algorithm: With both D_s and D_{chan} modeled as functions of E_s and B_s, the optimization problem is:

$$\min_{E_s,B_s}(D_s+D_{chan}),\quad 0<E_s<E_{TX},\ B_s>0 \tag{12}$$

The expressions for D_s, D_{chan} are (3) and (10). Because the optimal D_s and D_{chan} are both differentiable with respect to E_s and B_s, if (12) has a solution, the optimal B_s, E_s is a solution of equations in (13).

$$\frac{\partial D_{chan}}{\partial B_s} = -\frac{\partial D_s}{\partial B_s},\quad \frac{\partial D_{chan}}{\partial E_s} = -\frac{\partial D_s}{\partial E_s} \tag{13}$$

Substituting expressions for $\frac{\partial D_{chan}}{\partial E_s}$ and $\frac{\partial D_s}{\partial E_s}$ into the right equation in (13), we get

$$\begin{aligned} &const\cdot a\cdot e^{-a\cdot E_{t\_sb}^{(n)}+b}\cdot\frac{\sigma^2}{N_0} \\ &= a_2\cdot e^{-a_1\cdot B_s+b_1}\cdot e^{-a_2\cdot C_s+b_2}\cdot\left(\frac{1}{\beta\cdot\gamma\cdot C_s^{\gamma-1}}\right) \end{aligned} \tag{14}$$

We take the logarithm on both sides of (14) and rearrange the equation to obtain:

$$\begin{aligned} -a\cdot E_{t\_sb}^{(n)} + a_1\cdot B_s + a_2\cdot C_s + (\gamma-1)\cdot lnC_s \\ +[b-b_1-b_2+ln(\frac{a}{a_2}\cdot\beta\cdot\gamma\cdot\frac{\sigma^2}{N_0})] = 0 \end{aligned} \tag{15}$$

If we define

$$\begin{aligned} f \triangleq -a\cdot E_{t\_sb}^{(n)} + a_1\cdot B_s + A_2\cdot C_s + (\gamma-1)\cdot lnC_s \\ +[b-b_1-b_2+ln(\frac{a}{a_2}\cdot\beta\cdot\gamma\cdot\frac{\sigma^2}{N_0})] \end{aligned} \tag{16}$$

then $f=0$ is a condition for the optimal point. The expression shows that f is a function of $E_{t\_sb}^{(n)}$, B_s, and C_s. Dividing the left equation of (13) by the right, we get

$$\frac{(\frac{\partial D_{chan}}{\partial B_s})}{(\frac{\partial D_{chan}}{\partial E_s})} = \frac{(\frac{\partial D_s}{\partial B_s})}{(\frac{\partial D_s}{\partial E_s})} \tag{17}$$

Substituting the expressions for $\frac{\partial D_{chan}}{\partial B_s}$, $\frac{\partial D_s}{\partial B_s}$, $\frac{\partial D_{chan}}{\partial E_s}$, and $\frac{\partial D_s}{\partial E_s}$ into (17), we then take the logarithm of both sides and rearrange the equation. This yields

$$E_{t\_sb}^{(n)} = \frac{a_1}{a_2}\cdot\beta\cdot\gamma\cdot C_s^{\gamma-1}\cdot(1+e^{a_2\cdot C_s-b_2})\cdot\frac{\sigma^2}{N_0}-\frac{1}{a} \tag{18}$$

which shows that $E_{t\_sb}^{(n)}$ is a function of C_s. Note that

$$B_s = \frac{E_{TX}-E_s}{E_{t\_sb}} = \frac{E_{TX}-\beta\cdot C_s^{\gamma}}{E_{t\_sb}^{(n)}}\cdot\frac{\sigma^2}{N_0} \tag{19}$$

so B_s is also a function of C_s . Substituting (18) and (19) into (16), f becomes a univariate function of C_s. The optimal C_s is a solution of the equation

$$f(C_s) = 0 \tag{20}$$

With the parameter values obtained in fitting data of the sample videos, it is observed that $E_{t\_sb}^{(n)}$ satisfying (18) is an increasing function of C_s, which crosses 0 at some C_s value (denoted C_{s_0}) between 0 and 1. In $(C_{s\_0}, 1]$, $f(C_s)$ is observed to be a decreasing function of C_s, and there is a unique C_s at which (20) is satisfied. Based on these observations, a binary search can be applied to solve (20). The steps are listed in Table II.

TABLE II
Optimization Algorithm.

Step 1	Initialization: $C_s = 0.5$, step size $\Delta C_s = 0.25$
Step 2	Evaluate $f(C_s)$ using (18),(19),(16))
Step 3	If $E_{t\_sb}^{(n)} < 0$ or $f(C_s) > 0$, $C_s = C_s + \Delta C_s$; If $E_{t\_sb}^{(n)} > 0$ and$f(C_s) < 0$, $C_s = C_s - \Delta C_s$;
Step 4	If target precision of C_s is not yet achieved, return to Step 2
Step 5	If $E_{t\_sb}^{(n)} < 0$, $C_s = C_s + \Delta C_s$ (ensuring $C_s > C_{s\_0}$)
Step 6	Calculate the optimized $E_{t\_sb}^{(n)}$ using (18); Calculate optimized B_s using (19); Map the optimizaed B_s,C_s to QP and video configuration parameters; Obtain the optimal r according to Fig. 4; Calculate optimal E_b using (8).

Fix-E_{t_sb} algorithm: A simpler algorithm allocates sufficient transmission energy to ensure reliable transmission, so that end-to-end and source distortions are approximately equal. The $E_{t\_sb}^{(n)}$ which makes D_{chan} smaller than a target value is calculated. With $E_{t\_sb}^{(n)}$ fixed, E_s and B_s are constrained by (11) and the optimization is

$$\min_{C_s} D_s = e^{-a_1 \cdot B_s + b_1} \cdot (1 + e^{-a_2 \cdot C_S + b_2}) + D_{floor}$$
$$B_s = \frac{E_{TX} - \beta \cdot C_s^{\gamma}}{E_{t\_sb}^{(n)}} \cdot \frac{\sigma^2}{N_0} \quad (21)$$

The solution can be found by taking the derivative of D_s with respect to C_s (see B_s as a function of C_s), and searching for the C_s at which the derivative is 0.

VI. Simulation Results

We apply the proposed algorithms to a specific problem setup (see Table III) to measure end-to-end performance. Values of a and b are obtained by curve fitting, N_0 is the power spectral density of thermal noise at room temperature, and σ^2 reflects the path loss, calculated according to the Friis equation (distance=100m, carrier frequency= 2.4GHz). Curve-fitting parameters a_1, b_1 , a_2, b_2 , D_{floor} of the source model are obtained by fitting the data from encoding the sample video. The energy constraint E_{TX} varies in a range from one to seven times the maximum video encoding energy.

TABLE III
Problem setup for simulation.

Channel model parameters	Source model parameters
$E_{TX} = 0.01 \sim 0.07J$	$\beta = 0.01J, \gamma = 2.5$
$a = -0.0215, b = 292$	$a_1 = 8.53 \times 10^{-7}$, $b_1 = 4.52$
$N_0 = -174dBm$, $\sigma^2 = -83.05dB$	$a_2 = 9.83$, $b_2 = 1.07$
$c = 100$, $N_f = 30$	$D_{floor} = 4.87$

Two baseline algorithms are used. The least-complexity algorithm uses the configuration with the least complexity, and selects the QP to ensure that enough energy is allocated to achieve reliable transmission so that channel distortion can be ignored. The most-complexity algorithm is similar except that it uses the configuration with the highest complexity. These baselines represent a class of simple algorithms that pick an arbitrary configuration, and choose QP according to the energy requirement of transmission.

Fig. 5 shows PSNR plotted versus E_{TX}, obtained from experiments on test video sequence *Shadow*. The curve for the highest-complexity algorithm does not extend to small E_{TX} because there is not enough energy to encode with the most complex configuration. The two proposed algorithms outperform the baseline algorithms. The gap between the curves varies irregularly with E_{TX}, reaching 1dB at some E_{TX} values. The irregularity can be explained by the fact that the QP and the configuration are discrete, so it is impossible to find the QP and the configuration that give exactly the B_s and C_s produced by the algorithm. When the algorithm-produced B_s and C_s are mapped to a specific choice of QP and configuration, the algorithm's actual performance has some gap from the model. The curves for the two proposed algorithms are almost identical, so the fix-E_{t_sb} algorithm is as good as the binary-search algorithm.

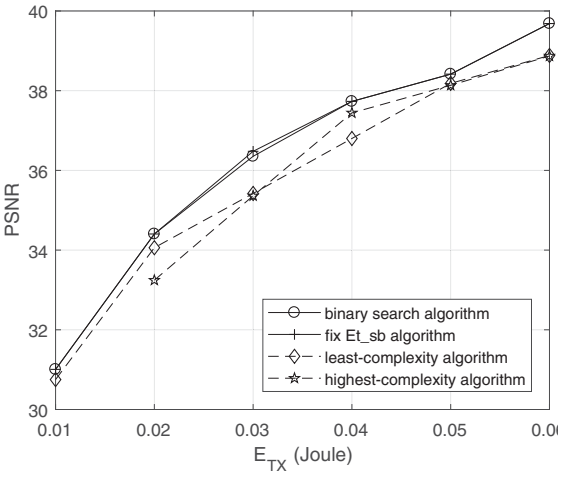

Fig. 5. Performance (PSNR) of algorithms as a function of the total energy constraint.

VII. Conclusion

We studied the joint optimization of video encoding and transmission, targeting wireless video sensors, which have stringent energy constraints. Under an energy constraint, we studied the minimization of the end-to-end distortion, by optimizing FEC code rate, energy allocation and number of source bits the video encoder generates. Comparing to baseline algorithms that employ no joint optimization, the proposed algorithms achieve a gain up to 1dB.

ACKNOWLEDGEMENT

This research was supported by NSF Grant IIS-1522125 and ARO Grant W911NF-14-1-0340.

References

[1] Y. Eisenberg, C. Luna, T. Pappas, R. Berry, and A. Katsaggelos, "Joint source coding and transmission power management for energy efcient wireless video communications", *IEEE Trans. Circuits Syst. Video Tech.*, vol. 12, pp.411-424, 2002.

[2] Z. He, W. Cheng, and X. Chen, "Energy minimization of portable video communication devices based on power-rate-distortion optimization", *IEEE Trans. Circuits Syst.*, vol. 18, no. 5, pp.596-608, 2008.

[3] 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 Tech.*, vol. 12, no. 6, pp. 511-523, 2002.

[4] S. Cui, A. Goldsmith, and A. Bahai, "Energy-constrained modulation optimization", *IEEE Trans. Wireless Commun.*, vol. 4, no. 5, pp. 2349-2360, 2005.

[5] Y. Liang and I. Ahmad, "Power and distortion optimization for pervasive video coding", *IEEE Trans. Circuits Syst. Video Tech.*, vol. 19, pp. 1436-1447, 2009.

[6] R. Zhang, S. L. Regunathan, and K. Rose, "Video Coding with Optimal Inter/Intra-Mode Switching for Packet Loss Resilience", *IEEE JSAC*, vol. 18, pp. 966-976, 2000.