

POWER AND CONTENT AWARE VIDEO ENCODING FOR VIDEO COMMUNICATION OVER WIRELESS NETWORKS

Yongfang Liang and Ishfaq Ahmad

Department of Computer Science and Engineering
University of Texas at Arlington, TX 76019

ABSTRACT

In the highly versatile next-generation wireless networks, one of fundamental problems is how to efficiently manage the power consumption while preserving high quality performance. This paper addresses power-and-content-aware video encoding in pervasive wireless/mobile environments. We formulate the problem of video quality maximization while preserving the battery power of the underlying device as an optimization problem. This is equivalent to minimizing both the power consumption and video distortion simultaneously. Based on a complexity scalable video coding architecture and the content aware adjustment, we propose a software based compression scheme. We call the proposed encoder active because it best uses its bits and power and tunes its parameters according accordingly. The preliminary results show that the encoder determines a suitable configuration according to the coding environment in order to save power and maintain well-perceived video quality.

1. INTRODUCTION

At the turn of the last century, we witnessed the genesis of the “wireless and pervasive computing era”, in which the focus of telecommunications and computing started to change, from traditional wired telephony-oriented services and infrastructures to data-based services, and from desktop workstations to hand-held, personal digital assistants, and smart sensors. In the near future, pervasive video applications are likely to provide new exciting services to users in a home or small office environment, such as mobile video conferencing, mobile video streaming, and sensor based video surveillance. Since the wireless networks are to accommodate battery-powered equipments with limited power supply, one fundamental problem is how to efficiently manage the power consumption while preserving well-perceived video presentation quality. This problem becomes even more critical in the highly versatile next-generation wireless networks.

Research efforts have been made to address the issue of limited power supply on mobile devices from two different ways. Power-aware design techniques [3], [6], [7] attempt to adjust source-coding parameters to maximize the performance under power dissipation constraints. At the other extreme, low-power design techniques try to lower the intensive complexity of video encoding, which in turn reduces the power consumption, with or without a desired performance target [1], [9], [10].

However, the reported work reviewed above could not meet the new requirements in the coming ubiquitous computing environments. We argue that video compression technologies in futuristic pervasive environments will have to provide high quality of service, which is measured from two perspectives: First, the video quality at a given bit rate must be optimized. Second, it should be able to efficiently utilize its power supply to prolong the operating time. In other words, the compression scheme should be active, it knows how to best use its bits and power and tune its parameters according to the content it is compressing. Such a concept is important in emerging applications. For example, the sensors should remain silent to save power unless triggered when the content becomes active [4].

This paper addresses the design of a power-and-content-aware video encoding system in the pervasive wireless/mobile environments. In this work, our focus is on techniques for efficiently utilizing the energy supply while preserving desirable video quality. We consider not only the video encoding architecture, but power and quality control schemes as well. We formula the power-and-content-aware video encoding problem and discuss the approach to solve this problem. To utilize the power efficiently, we propose a power scalable video encoding architecture with embedded parametric algorithms and content aware power-efficient feedback control schemes, which focus on how to adjust the video distortion parameter as demanded by the environment. Based on the proposed architecture and control schemes, we develop a power-and-content-aware video encoding system that is able to prolong the battery lifetime while gaining desirable video quality.

The remainder of this paper is organized as follows. In Section 2, we formulate the problem of power-distortion-optimized video coding. The constraint-oriented approach to solve this problem is also presented. Section 3 introduces the complexity scalable video encoder, followed by the content aware adjustment strategy to further save the power energy. Section 4 summarizes the power-and-content-aware video coding system and its adjustment strategy. Section 5 presents the simulation results. Concluding remarks are provided in Section 6.

2. POWER-DISTORTION-OPTIMIZED VIDEO CODING ANALYSIS

In this section, we formalize the optimization problem of video coding in wireless communication environments. The constraint-oriented approach is applied to solve this problem.

2.1. Power-distortion-optimized problem

To formulate the research problem and present the main concept, we make a few assumptions about the underlying hardware platform:

1. The power supply of the system is attainable through low-level circuit design. Whenever wanted, the current power supply is reported to the high-level modules.

2. A mapping function to translate the computational complexity of the video coding system into corresponding power consumption is available, and vice visa.

The first assumption can be worked around by using several off-the-shelf techniques proposed in the literature, such as estimating the remaining power using microcontroller [2], or the “state-of-charge” techniques [5]. With the second assumption, the power consumption can be tuned by the complexity parameters. Therefore, a complexity scalable video encoder can be eventually translated into an energy consumption scalable system.

Let $C(R,X)$, $D(R,X)$ denote the computational complexity and video distortion respectively, with complexity parameter set X at coding bit rate R . Let Φ be the mapping function. The power consumption of the complexity scalable system is given by:

$$P(R,X) = \Phi(C(R,X)). \quad (1)$$

In video coding applications in wireless sensor networks, our objective is to minimize not only the distortion but also the power consumption as well. Thus, for the scalable video encoder, the goal is to find a complexity parameter set that minimizes the power consumption and video distortion. Mathematically, we have:

$$\begin{aligned} \min_{X \in \zeta^N} F(R,X) &= [P(R,X), D(R,X)]^T, \\ \text{subject to } P(R,X) &\leq P_c \end{aligned} \quad (2)$$

where ζ^N is the space of N complexity parameters and P_c is the current power constraint for video coding. In general, (2) is a multiple-objective optimization (MOO) problem with nonlinear constraint.

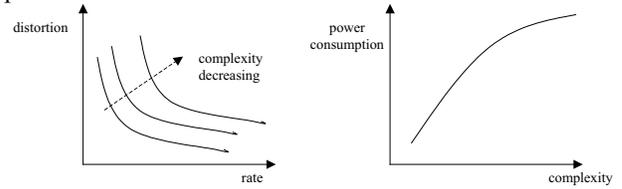


Fig. 1: The relations of rate-distortion-complexity and power-complexity.

Generally, a more efficient algorithm has higher complexity, which in turn results in smaller distortion and higher power consumption on processing. In contrast, a relatively simple algorithm has lower power consumption at the expense of bigger distortion. Fig. 1 shows the rate-distortion-complexity and power-complexity relations. The lower the coding bit rate, the larger the distortion is. At a given coding bit rate, higher complexity results in lower distortion but with higher power consumption. Thus, the objective functions in the above power-distortion-optimized problem are incommensurate and in conflict with one another with respect to their minimum goals. There is no single optimal solution to (2). We need to search for a best tradeoff between minimal distortion and minimal power consumption to ensure a satisfactory design.

2.2. Constraint oriented approach

Various approaches have been proposed to solve a MOO problem by transforming it into a substitute problem. In this work, we apply the constraint-oriented approach. One objective function is used as the main objective and the other is treated as the secondary objective. Considering the fact that the human being can tolerate some video distortion, in power-distortion-optimized video coding, we can exploit this property without having to make it obvious or without compromising on the obvious visual quality in order to save power. Therefore, we treat minimal power consumption as the main objective function and minimal distortion as the secondary objective function. Theoretically, we can substitute the above MOO problem with:

$$\begin{aligned} \min_{X \in \zeta^N} P(R,X) \\ \text{subject to } P(R,X) \leq P_c, D(R,X) \leq D'' \end{aligned} \quad (3)$$

where D'' is the upper bound of video distortion that can be determined by the particular applications.

It is worth to mention that the upper bound of the cost function influences the results of the substitute problem. In practice, the constraint of the upper bound is very subjective and it is application dependent. In order to

achieve the best performance, we try to obtain the minimum distortion while taking the power consumption into consideration. The upper bound distortion is achieved by solving the distortion-optimized problem, given by:

$$D^u = \min_{X \in \mathcal{E}^N} D(R, X), \quad (4)$$

subject to $P(R, X) \leq P_c$

Now, to obtain the solution of problem (2), we solve (3) using the desirable distortion upper bound from (4). In general, (3) and (4) are nonlinear programming (NLP) problems. Both the objective and the constraint functions are nonlinear. Optimization methods such as Lagrange multiplier and penalty function methods can be used to solve the above optimization problems.

3. COMPLEXITY SCALABLE AND CONTENT AWARE VIDEO CODING

In the previous section, we have analyzed the power-distortion-optimized video coding problem. In this scenario, to achieve flexible management and control of power consumption, we need to develop a video encoding architecture, which is fully scalable in power consumption. Moreover, to provide adaptability and self-adjustment, the encoder should be content aware.

3.1. Complexity scalability

Typical video encoders, such as MPEG-1, MPEG-2, H.263, MPEG-4, and the newly developed H.264/AVC, employ a hybrid motion compensated DCT (discrete cosine transform) encoding scheme and include some general encoding modules such as ME (motion estimation), DCT/IDCT, quantization/dequantization, reconstruction, and entropy coding. It implies that these modules are the candidates for complexity management.

3.1.1. Power consumption analysis

Experimental results show that for most video coding standards, the ME and the PRECODING modules, which include DCT, IDCT, quantization, dequantization and reconstruction modules, consume most of the CPU cycles, while the entropy coding module only uses a relatively small amount of total CPU cycles, especially at low coding bit rates. Meanwhile, the frame rate affects the computational complexity dramatically. To show the concept of power content aware video coding, in this work, we use the ME module, PRECODING modules and the frame rate as the complexity control modules. The same concept can be applied to other video coding systems, such as the latest H.264/AVC in which the CABAC (context-based Adaptive Arithmetic Coding) entropy coding and the deblock operation also consume a lot of computational power and need to be controlled. In the following, we denote x , y , and z the normalized

complexity parameters of the ME module, the PRECODING modules and the frame rate respectively, given by:

$$x = \lambda_{ME} / \lambda_{ME}^{max}, \quad y = \lambda_{PRE} / M, \quad z = f / f_{max}, \quad (5)$$

where M is the number of MBs in one frame, λ_{ME}^{max} is the maximum value of λ_{ME} , and f_{max} is the maximum frame rate.

For the ME module, the computational complexity is determined by the number of SAD (sum of absolute difference) computations (denoted by λ_{ME}) associated with each frame. We dynamically allocate the available number of SAD computations throughout the frame among the MBs according to their motion characteristics. We refer this scheme as λ_{ME} scalability. For the PRECODING modules, after motion estimation and compensation, the SAD values of the M MBs are sorted in an ascending order and the first $M - \lambda_{PRE}$ MBs are forced to be AZMBs (all zero MB) in which the PRECODING operations are skipped. This scheme is referred as λ_{PRE} scalability.

By introducing complexity parameters to control the major power demanding encoding modules, the video encoder is fully complexity scalable. The computational complexity for each coding frame is proportional to the number of SAD, the number of non-zero MB, and the coding rate. Thus, the power consumption is given by:

$$C(R, x, y, z) = z(C_1x + C_2y + C_3R), \quad (6)$$

where C_1 , C_2 , C_3 are the complexity ratios for the ME, PRECODING, and entropy coding respectively. They can be obtained either by theoretical cycle estimation or from simulation statistics.

3.1.2. Rate-distortion model

The derivation of the rate-distortion model of the system can be found in our previous work [6]. In this paper, because of its simplicity, we use the average SAD instead of SSD (sum of square difference) to estimate the average MB variance. By using similar analysis, the rate-distortion behavior can be modeled as:

$$D(R, X) = D(R, x, y, z) = D_0 + D_1, \quad (7)$$

where

$$D_0 = 4 \times (1 - z)^2 (\beta_0 + \beta_1)^2 \quad (8)$$

$$D_1 = 8 \times (2z - z^2) (\beta_0 + \beta_1^{-\beta_2 x})^2 \left(\frac{1}{2} (1 - y)^2 + y(1 + a_0 y) \times 2^{-2\gamma R / y} \right)$$

In (8), β_0 , β_1 , and β_2 are the model parameters of the ME module, determined by the characteristics of the video scene. γ and a_0 are the model constants of the PRECODING modules.

As far as the video quality is concerned, the proposed scheme maintains a reasonable good performance. It is well known that the moving objects in the video scene contribute most to the overall visual quality and they often have relatively large SAD values after ME. In our schemes, the λ_{ME} scalability distributes the available number of SAD computations among the MBs and allocates more SAD computations to the moving objects.

Moreover, the PRECODING operations are applied to the MBs with large SAD values in the λ_{PRE} scalability. Thus, more computation will be allocated to the moving objects, which results in improved overall video quality.

3.2. Content awareness

When solving the power-distortion-optimized MOO problem, the upper bound of desirable video distortion plays an important role in power saving. As the upper bound increases, the video encoder can run at a lower complexity, which results in smaller power consumption from the power saving perspective. Therefore, we introduce a tolerance parameter for more power saving purpose. Let δ be the video distortion tolerance, (3) becomes:

$$\begin{aligned} \min_{X \in \mathcal{X}} P(R, X) \\ \text{subject to } P(R, X) \leq P_c, D(R, X) \leq D'' + \delta \end{aligned} \quad (9)$$

Since the human being can tolerate some video distortion, we can expect that the video quality will not degrade much when we slightly relax the upper bound. The degrading in perceived video quality is negligible with respect to the improvement in power saving. Moreover, the human eyes are more sensitive to fast motion than low motion. For low motion sequences, we can have bigger tolerance. Prediction of the motion characteristic enables further performance improvement in terms of more power energy saving.

To make the tolerance parameter adaptive to the coding sequences, it is important for the video encoder to know the coding content. We must remember that the encoder has limited memory and limited computational power and thus we need fast, simple and yet efficient methods to understand and store the meaning of content.

We have proposed a novel hierarchical framework to fully retrieve the subtle motion information of a video sequence. We showed the effectiveness of the concept on developing a fast motion estimation algorithm in [8]. In this framework, the motion intensity is determined at three levels: block level, frame level, and sequence level. The sequence level motion intensity represents the motion characteristic of the video sequence. Notice that the sequence level motion intensity is derived from the temporal motion distribution, and it will change when the coding environment changes, for example when the capturing switches from the static stage to the acting actors. In this work, we use it to classify the video sequences.

According to the sequence level motion intensity, video sequences are classified in three classes: a) low motion; b) medium motion; c) high motion. Based on the classification, we change the tolerance parameter in (9) accordingly. Currently, we choose the perceptible tolerance as 5%, 8%, and 10% of D'' for low, medium, and

high motion sequence respectively. Subjective experiments show that the people will not notice the difference of the perceived video quality with these small changes. As a result, the adjustment of the tolerance parameter is adaptive to the content it is compressing.

As the extra computational complexity is concerned, the beauty of the proposed technique is that it requires no extra expensive computation as the required information is extracted from the previously calculated motion vectors, which is already available after motion estimation.

4. POWER-AND-CONTENT-AWARE VIDEO CODING

Having formulated the research problem and introduced both the complexity scalability and the dynamic adjustment of the video distortion tolerance, now we propose a power and content aware video coding system in wireless sensor networks, minimizing the distortion and power consumption.

Fig. 2 shows the architecture of the power and content aware video coding system. The complexity scalable schemes based on λ_{ME} scalability and λ_{PRE} scalability are embedded into this architecture. By choosing appropriate complexity parameters, the system can be tuned to the optimal operating state. The system analyzes the coding content based on the hierarchical motion intensity framework and adjusts the tolerance of the video distortion. The energy-monitoring module provides the current power constraint. The bit allocation module determines the number of bits to encode the next frame or group of frames according to the available bit rate.

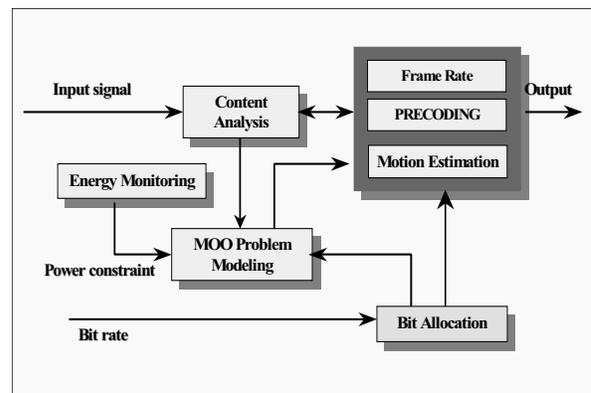


Fig. 2: The power and content aware encoding system.

Using the techniques discussed above, we are able to achieve the minimal power consumption while preserving the video quality. The proposed power and content aware video encoding system operates as follows:

Step 1: Power Monitoring. The power consumption of the system is monitored to get the current available power supply.

Step 2: Model Parameter Estimation. The system model parameters in (8) are estimated from the statistics of previous frames.

Step 3: Optimization and Complexity Control. Solve the MOO problem to get the complexity parameters (x,y,z) in a specific time period to update the system parameters. Set the corresponding frame rate; perform λ_{ME} complexity and λ_{PRE} complexity control.

Step 4: Dynamic Parameter Adjustment. For each encoding frame, after motion estimation and compensation, we collect the motion information to analyze the coding content and adjust the distortion tolerance parameter if needed.

5. SIMULATION RESULTS

To evaluate the effectiveness of the proposed video encoding system, we implement it into the public domain H.263+ encoder. The purpose of this simulation is to demonstrate and validate the proposed power-distortion-optimized scheme for video coding. To simplify the measurement of power consumption, we use a linear mapping function to translate the computational complexity into power consumption. For a given power constraint, we convert it into the corresponding complexity constraint through the mapping function. The translated complexity is then used as the input of the complexity scalable video encoder. In this simulation, the power consumption is normalized to 1. λ_{ME}^{max} is 50 and f_{max} is 30fps. The video distortion is measured by the mean square error (MSE).

For the "Foreman" QCIF sequence at 128kbps/15fps under different given power supply, Fig. 3 and Fig. 4 show the complexity parameters $\{x,y,z\}$ as functions of the percentage of power consumption without and with the content aware adjustment. We can see that with the adjustment, the encoder can lower its computational complexity quickly to minimize the power consumption and distortion. Simulation over other test video sequences yields similar results. Fig. 6 illustrates the corresponding power consumption and video distortion with/without the dynamic content aware tolerance adjustment. As we can see, after applying the adjustment, the difference in video quality is negligible. However, from the power saving perspective, one can observe that the proposed adjustment consumes much less power and therefore it is more power efficient. From the subjective video quality point of view, we can see little difference in the reconstructed video scenes, as shown in Fig. 5. Fig. 7 and Fig. 8 give the results for other two different video sequences respectively. We can see that the power-and-content-

aware control is more power efficient with minor degraded video quality.

6. CONCLUDING REMARKS

In this paper, we proposed a new paradigm of video encoding: power and content aware video coding. As opposed to the traditional compression schemes, the proposed encoding scheme knows how to best use its bits and power in different coding environments. We showed that the proposed power-distortion-optimized video coding model can be considered as a multiple objective optimization problem. Based on a complexity scalable video coding system and the content aware adjustment, the encoder adjusts its configuration and the complexity parameters to save processing power and maintain well-perceived video quality. Simulation results demonstrated that the proposed scheme is aware of the power of the underlying device as well as the video content, and achieves significant power saving within a tolerable distortion range. Such a feature is particularly desirable for video communication applications in wireless sensor network.

7. REFERENCES

- [1] S. M. Akramullah, I. Ahmad, and M. L. Liou, "Optimization of H.263 video encoding using a single processor computer: performance tradeoffs and benchmarking," *IEEE Trans. On Circuits and System for Video Technology*, vol. 11, pp. 901 - 915, August 2001.
- [2] L. Bowen, R. Zarr and S. Denton, "A Microcontroller Controlled Battery Fuel Gauge and Charger," *Proceedings of the Ninth Annual Battery Conference on Applications and Advances*, 11-13 Jan. 1994.
- [3] Wayne Burleson, Prashant Jain, Subramanian Venkatraman, "Dynamically Parameterized Architectures for Power-Aware Video Coding: Motion Estimation and DCT," *Workshop on Digital and Computational Video 2001*, 4-12.
- [4] J. Carle and D. Simplot-Ryl, "Energy-Efficient Area Monitoring for Sensor Networks," *IEEE Computer*, pp. 40-46, Feb. 2004.
- [5] O. Caumont, P. Le Moigne, C. Rombaut, X. Muneret and P. Lenain, "Energy Gauge for Lead-Acid Batteries in Electric Vehicles," *IEEE Transactions on Energy Conversion*, Volume: 15, Issue: 3, Sept. 2000.
- [6] Yongfang Liang, Zhihai He, and Ishfaq Ahmad, "Analysis and Design of Power Constrained Video Encoder," *IEEE 6th CAS Symposium on Emerging Technologies: Frontiers of Mobile and Wireless Communication (MWC'04)*, Shanghai, China, May 31-June 2, 2004.
- [7] T. Lan and A.H. Tewfik, "Power optimized mode selection for H.263 video coding and wireless communications,"

[8] Yongfang Liang, Ishfaq Ahmad, Jiancong Luo, Yu Sun, "Fast Motion Estimation Using Hierarchical Motion Intensity Structure," in *ICME 2004*, Taipei, Taiwan, June 2004, in Press.

[9] X. Lu, Y. Wang, E. Erkip, "Power Efficient H.263 Video Transmission over Wireless Channels," *Proc. Of IEEE ICIP*, Rochester, New York, September 2002.

[10] Qian Zhang, Zhu Ji, Wenwu Zhu, and Ya-Qin Zhang, "Power-minimized bit allocation for video communication over wireless channels," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol.12, No. 6, P398 - 410, June 2002.

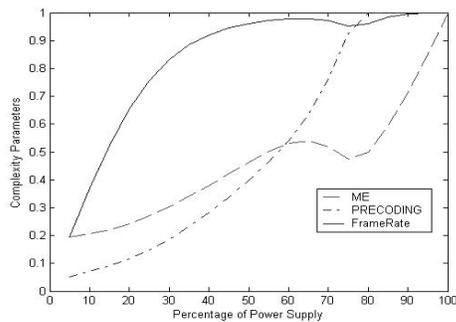


Fig. 3: Complexity parameters for "Foreman" QCIF sequence 128kbps at 15fps.

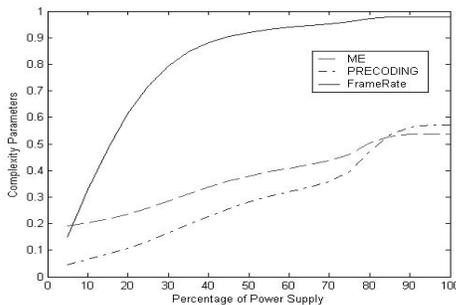


Fig. 4: Complexity parameters after content aware adjustment for "Foreman" QCIF sequence 128kbps at 15fps.



Fig. 5: Comparisons of video quality: (a) original result; (b) optimized result.

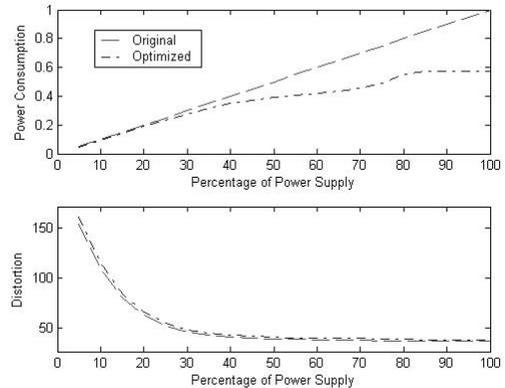


Fig. 6: Power consumption and distortion comparisons for "Foreman" QCIF sequence 128kbps at 15fps.

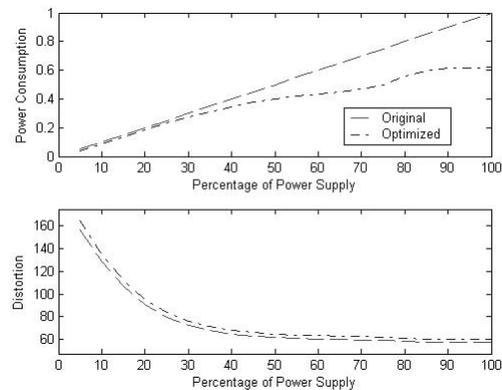


Fig. 7: Power consumption and distortion comparisons for "Carphone" QCIF sequence 128kbps at 15fps.

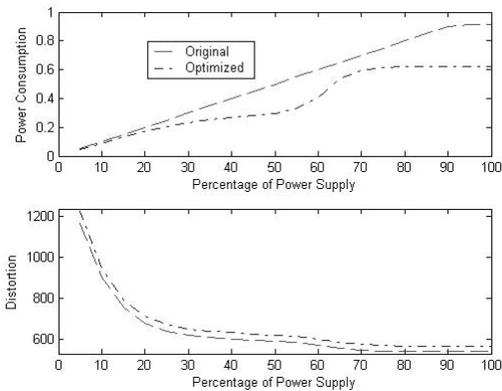


Fig. 8: Power consumption and distortion comparisons for "Stefan" QCIF sequence 64kbps at 15fps.