COMPUTATION AND TRANSMISSION ENERGY MODELING THROUGH PROFILING
FOR MPEG4 VIDEO TRANSMISSION

Amol Bhatkar and R. Chandramouli*  
Dept. of Electrical and Computer Engineering,  
Stevens Institute of Technology

N. Vijaykrishnan and M .J. Irwin  
Dept. of Computer Science and Engineering,  
Penn State University

ABSTRACT
Video Communications using mobile wireless devices is a challenging task due to the limited capacity of batteries. Some of the key technologies that affect battery life are source compression, channel error control coding and radio transmission. In this paper, we propose several experimental profiling based mathematical models for energy consumption of MPEG4 video transmission. These models capture both the computational and transmission energy as a function of several MPEG4 parameters. Results presented in this paper show that the proposed models are fairly accurate.

1. INTRODUCTION
Energy efficient wireless video transmission is a key challenge in mobile networking. Due to the complex coding algorithms, energy consumed when processing has become a significant component of battery drain. Video compression typically exploits the redundancy between frames to achieve higher compression rates. MPEG4 use the motion compensated picture difference technique to achieve high degrees of compression at acceptable levels of picture quality. Also more the source is compressed, less is the source rate, which indicates a reduction in the total transmission energy. Note that the transmission energy is also dependent on the distance between the transmitter and receiver and, the underlying channel conditions. Higher compression also results in a higher distortion level as compression ratio increases. The goal of an energy efficient mobile wireless video communications is to minimize the total energy dissipation, while keeping the end-to-end distortion constant. Energy models can help in such situations by providing before-hand energy consumption values which can be used for achieving the above objective. Energy or power optimization problems for wireless networking have been studied by many researchers in a variety of ways. In particular, [1], [2], [3], [4], [5], [6], [7] deal with energy efficient video/image transmission techniques. In our work, we look at energy models for the MPEG4 video codec.

The goal of this paper is to build energy consumption models for both computation and transmission energy for MPEG4 [8] video transmission. Towards this objective we use profiling and mathematical regression techniques. Even though we have built energy models for several MPEG4 modes and parameters using different video sequences, here we present only a subset of these results due to space constraints. The impact of motion-estimation search range and intra rate on computational energy efficiency and the resulting tradeoff in transmission energy efficiency are given in detail. We present a brief overview of MPEG4 in Section 2, the experimental setup in Section 3 and energy modeling in Section 4. Conclusions are given in Section 5.

2. BRIEF OVERVIEW OF MPEG4

MPEG4 is an ISO/IEC standard developed by Motion Picture Experts Group(MPEG). The standard is expected to be a predominant encoding for wireless and Internet video, and offers significant advantages over current formats due to its ability to code video sequences on an object-by-object basis. It permits compression of rectangular sized image sequences at varying levels of input formats, frame rates and bit rates while providing comparable quality. While spatial redundancy is exploited via transform coding, block-based motion estimation and compensation techniques are employed for exploiting temporal redundancy. I-frames are intra-coded frames, coded independently of other frames, so they act as reference frames. P-frames are coded predictively from the closest previous reference frame, and B-frames are coded bi-directionally from the preceding and succeeding reference frame. Dependent P and B frames are coded using motion estimation, which includes a process called as block matching and motion compensation. In block matching, pixel blocks of a dependent frame, called macroblocks are matched against macroblocks in the reference frame to find the closest possible match for that block. The amount by which these matching blocks are displaced with respect to each other is encoded as motion vector. The motion residual is then encoded into the bitstream along with the motion vectors. Now we consider the motion vector search range. Since the motion vector search range represents the macroblock dimensions, as it goes on increasing the computational energy increases due to increase in search area. The typical optimal search range varies between sequences depending on the motion intensity factor. It is seen that if the search range becomes larger than 99 % statistics of the sequence the coding quality decreases and source rate increases. This can be attributed to the fact that as search range increases, large motion vectors are been selected, and more bits are used to code the motion-vector differences. These additional bits are however not sufficiently compensated for by corresponding savings in coding smaller residual errors. As we can see, a number of different parameters affect the computational energy consumption. We present our experiments based profiling approach for energy modeling in the next two sections.
3. EXPERIMENTAL SETUP FOR ENERGY PROFILING

We use the Microsoft MPEG4 visual reference software [9], which is an implementation of MPEG4 video tools as specified by the MPEG4 video standard (ISO/IEC 14496-2). The testbed is a Sony Vaio laptop with a 700 MHz P-III processor, 128 MB RAM, running Red Hat Linux 2.4.8 chosen for its open source nature. The power consumed by the CPU in running the MPEG4 coder is measured as a function of input power supply to the Laptop. A separate DC power supply is given to the Laptop to permit measurements. The battery of the laptop is removed for accuracy in measurements. The current measurements are gathered using Labview from the GPIB interface of the power supply (see, Fig. 1). To eliminate effects of the programs running in background, the current consumption $I_{idle}$ is first tested when no other tasks are running. The difference in currents when the coder is running and $I_{idle}$ is taken as the actual current consumption. In the experiments, since voltage variation is seen to be extremely small (measured at less than 0.25%) we use a constant value. To calculate the time taken to encode a video sequence we use oprofile [10], a system profiler for Linux. The profiler periodically sends interrupts and records information which includes the program counter (PC) and process identifier (PID) of currently executing processes. From this we compute the total time taken by each process.

4. ENERGY MODELING

In this section we describe our experimental methodology and associated mathematical models for energy consumption. Here we present results only for Foreman and Mother-daughter QCIF video sequences encoded at 10 fps using 400 and 961 frames respectively. These sequences are selected because of their different characteristics in motion and spatial detail. The motion intensity change $(a)$ [9] of these sequences were found to be 17.7% and 49 % respectively. This implies that Foreman sequence has a high motion intensity than the Mother-daughter sequence. For our experiments we have not used any rate control algorithm and skipping of macroblocks is also disabled, i.e., it prevents the encoder from skipping macroblocks in P-frames that do not change between frames. Due to this the time taken to encode individual frames becomes reasonably same across sequences for same intra rate. Thus the models using profiled data, becomes independent of content and can be used to predict MPEG4 encoder’s energy consumption for other QCIF video sequences. The models that we present are based on mathematical regression methods applied to data collected from the profiling experiments.

The base model was computed using profiled data for mother-daughter sequence when the intra coded macroblock percentage or intra rate $(\beta)$ was equal to 25%. This indicated a quadratic variation for the computational energy w.r.t. motion vector search range. But we note that other parameters such as the number of P and B frames, number of macroblocks in one frame etc. also play a crucial role in determining the energy consumption. Therefore, we propose the following model:

$$E(W) = N_J N_{MB} [(a + bW + cW^2) + \Theta_P(\beta, W) + \Theta_B(\beta, W)],$$  
(1)

where $E(W)$ is the computational Energy, $W$ is the Motion Vector Search Range, $N_J$ is the number of frames in the video sequence, $N_{MB}$ is the number of macroblocks in one frame, $\Theta_P(\beta, W)$ is a number of P frames dependent parameter, $\Theta_B(\beta, W)$ is a number of B frames dependent parameter and $a$, $b$, $c$ are constants. Eq. (1) has been optimized to minimize the magnitude of error across other QCIF sequences. Hence some amount of error is seen to exist (see Figs. 2 and 3) for even the Mother-daughter sequence. $\Theta_P(\beta, W)$ and $\Theta_B(\beta, W)$ account for the change in number of P frames and B frames respectively. Both are calculated in terms of intra rate. The intra rate can be varied by changing the number of P-frames and B-frames. For a particular search range as the intra rate decreases the computational energy increases as it takes more time to encode the inter frames. So as the number of P frames and B frames increases the computational energy also increases. We express $\Theta_P(\beta, W)$ as:

$$\Theta_P(\beta, W) = G_P(\beta) + H(W),$$  
(2)

where $G_P(\beta)$ is an intra rate only dependent function and $H(W)$ depends only on the search window size. Eq. (2) has been decoupled this way because the intra rate and the search range are not dependent on each other. From the data we observed that $G_P(\beta)$ is approximately linear with respect to the intra rate. Therefore, we have,

$$G_P(\beta) = G_P(\beta_0) + m(\beta - \beta_0),$$  
(3)

where $\beta_0$ is fixed intra rate and $m$ is the slope. The parameter $H(W)$ is a correction factor computed as a function of the motion search window size. This correction factor aids in producing a more accurate value for $\Theta_P(\beta, W)$. We also observe that $\Theta_P(\beta, W)$ versus the intra rate is a reciprocal quadratic equation. This is due to the following. It is known that B frames take the maximum computation for encoding as they are predicted bi-directionally. Since we are representing this in terms of intra rate, as number of B frame increases intra rate decreases and correspondingly computational energy increases. This along with the arguments presented previously to justify the form of Eq. (2) gives,

$$\Theta_B(\beta, W) = G_B(\beta) + H(W)$$  
(4)

and

$$G_B(\beta) = \frac{1}{x + y\beta + z\beta^2},$$  
(5)

where $x$, $y$, $z$ are constants. Fig. 2 shows the computational energy models (as given in Eq. (1)) for the Mother-daughter and Foreman sequences when intra rates are 25%, 16% and 12%. It is observed that as intra rate decreases the computational energy increases. This is because intra rate decreases due to increase in P frames and B frames respectively. Both are calculated in terms of intra rate. The intra rate can be varied by changing the number of P-frames and B-frames. For a particular search range as the intra rate decreases the computational energy increases as it takes more time to encode the inter frames. So as the number of P frames and B frames increases the computational energy also increases. We express $\Theta_P(\beta, W)$ as:

$$\Theta_P(\beta, W) = G_P(\beta) + H(W),$$  
(2)

where $G_P(\beta)$ is an intra rate only dependent function and $H(W)$ depends only on the search window size. Eq. (2) has been decoupled this way because the intra rate and the search range are not dependent on each other. From the data we observed that $G_P(\beta)$ is approximately linear with respect to the intra rate. Therefore, we have,

$$G_P(\beta) = G_P(\beta_0) + m(\beta - \beta_0),$$  
(3)

where $\beta_0$ is fixed intra rate and $m$ is the slope. The parameter $H(W)$ is a correction factor computed as a function of the motion search window size. This correction factor aids in producing a more accurate value for $\Theta_P(\beta, W)$. We also observe that $\Theta_P(\beta, W)$ versus the intra rate is a reciprocal quadratic equation. This is due to the following. It is known that B frames take the maximum computation for encoding as they are predicted bi-directionally. Since we are representing this in terms of intra rate, as number of B frame increases intra rate decreases and correspondingly computational energy increases. This along with the arguments presented previously to justify the form of Eq. (2) gives,

$$\Theta_B(\beta, W) = G_B(\beta) + H(W)$$  
(4)

and

$$G_B(\beta) = \frac{1}{x + y\beta + z\beta^2},$$  
(5)

where $x$, $y$, $z$ are constants. Fig. 2 shows the computational energy models (as given in Eq. (1)) for the Mother-daughter and Foreman sequences when intra rates are 25%, 16% and 12%. It is observed that as intra rate decreases the computational energy increases. This is because intra rate decreases due to increase in P frames and B frames respectively. Both are calculated in terms of intra rate. The intra rate can be varied by changing the number of P-frames and B-frames. For a particular search range as the intra rate decreases the computational energy increases as it takes more time to encode the inter frames. So as the number of P frames and B frames increases the computational energy also increases. We express $\Theta_P(\beta, W)$ as:

$$\Theta_P(\beta, W) = G_P(\beta) + H(W),$$  
(2)

where $G_P(\beta)$ is an intra rate only dependent function and $H(W)$ depends only on the search window size. Eq. (2) has been decoupled this way because the intra rate and the search range are not dependent on each other. From the data we observed that $G_P(\beta)$ is approximately linear with respect to the intra rate. Therefore, we have,

$$G_P(\beta) = G_P(\beta_0) + m(\beta - \beta_0),$$  
(3)

where $\beta_0$ is fixed intra rate and $m$ is the slope. The parameter $H(W)$ is a correction factor computed as a function of the motion search window size. This correction factor aids in producing a more accurate value for $\Theta_P(\beta, W)$. We also observe that $\Theta_P(\beta, W)$ versus the intra rate is a reciprocal quadratic equation. This is due to the following. It is known that B frames take the maximum computation for encoding as they are predicted bi-directionally. Since we are representing this in terms of intra rate, as number of B frame increases intra rate decreases and correspondingly computational energy increases. This along with the arguments presented previously to justify the form of Eq. (2) gives,
Now we study the effect of the intra rates on the transmission energy. It is known that for an uncoded BPSK system in additive zero mean white Gaussian noise channel [11],

\[ P_b = Q \left( \sqrt{\frac{2E_b}{N_0}} \right) \]  

(6)

where \( Q(\cdot) \) is the complementary error function, \( P_b \) is the bit error probability, \( E_b \) is the energy per bit and \( N_0 \) is the noise power spectral density. Thus,

\[ E_b = \frac{N_0}{2} \text{erf}^{-1}(1 - P_b) \]  

(7)

Assuming the single-sided power spectral density \( N_0 = 2 \times 10^{-5} \) W/Hz, the total transmission energy is calculated as \( E_b \) times the compressed video file size. The test sequence used is Foreman and the channel is assumed to be random. Fig. 4 shows the plots for Eqs. (3) and (7) obtained by varying the number of B frames. The total transmission energy is shown for different \( P_b \) values. Clearly we see that as the \( P_b \) constraint is relaxed the total transmission energy decreases. Also when the intra rate increases the compression ratio decreases, resulting in larger bit rates. Therefore we observe an increase in the total transmission energy. From the same figure, we also note that with an increase in intra rate the total computational energy decreases. This is because the number of B frames decreases thereby resulting in fewer computations. One interesting observation for this figure is that the rate of decrease of the computational energy is higher than the rate of increase of corresponding transmission energy. Similar results are seen for Eq. (3) and but are not shown here. Thus given a total energy budget the proposed models can be used to select appropriate values for coding and transmission parameters to meet the constraint.

Fig. 2. Energy model for profiled video: Mother-Daughter (top) and Foreman (bottom). The models are for different number of P frames.

Fig. 3. Energy Model for profiled video: Mother-Daughter (top) and Foreman (bottom). The models are for different number of B frames.
Intra rate
Total Computational Energy (Joules)

Intra rate
Total transmission energy (Joules)

Fig. 4. Comparison of total computational energy and total transmission energy w.r.t intra rate. The intra rates are obtained by varying the number of B frames and the total transmission energy is given for different BER ($P_b$).

5. CONCLUSION

We have derived profiling based energy consumption models for MPEG4 video coding and transmission. It is seen that the computational energy is a second order quadratic equation of the search range. The models vary linearly with the number of P frames, while they represent a reciprocal quadratic equation with the number of B frames. These models have been found to be reasonably accurate in predicting the computational energy consumption for a chosen motion vector search range. A desirable feature of the proposed models are their limited dependence on the content of the sequence. Interestingly we also observe that when the intra rate is increased, the rate of decrease of computational energy is higher than the rate of increase of corresponding transmission energy. Using the proposed models several MPEG4 video coding parameters can be easily chosen to satisfy a total energy constraint.

6. REFERENCES