ENERGY CONSUMPTION ANALYSIS AND MODELLING OF A H.264/AVC INTRA-ONLY BASED ENCODER DEDICATED TO WVSNS

Othmane Alaoui-Fdili[1,2]*,Youssef Fakhri*[1,3]*,Patrick Corlay*[2]*,François-Xavier Coudoux*[2]*,Driss Aboutajdine*[1]

[1]LRIT-CNRST URAC29, Université Mohammed V-Agdal, Rabat, Morocco;
[2]IEMN UMR 8520, Department OAE, UVHC, Valenciennes, France;
[3] LARIT équipe réseaux et Télécommunications, Université IbnTofail, Kénitra, Morocco.

ABSTRACT

This paper proposes a model to predict the energy consumption of a H.264/AVC intra-only based encoder designed for the wireless video sensor networks (WVSNs). Such model is of great interest in an energy-constrained context like the WVSN, where the video streams are processed prior to transmission. In fact, accounting for both processing's and transmission's consumed energies allows the optimization of the global energy consumption. The proposed model predicts the processing's energy consumption based on the considered Quantization Parameter (QP) and the Frame Rate (FR) values. The generic form of this model enables it to predict the energy consumption of any H.264/AVC intra-only based encoder. Simulation results demonstrate the accuracy of the proposed model, that is validated under different resolutions, QP and FR, with an average prediction error of 4%.

***Index Terms*—** Energy consumption modelling, video compression, H.264/AVC, wireless video sensor network

1. INTRODUCTION

A wireless sensor network is a self-organized system which is composed of many geographically distributed sensor nodes that collaborate to capture and process data, then route it to destination called the sink, via several short range wireless transmissions. In a wireless video sensor network (WVSN), the nodes are able to capture video streams thanks to the equipped low-cost CMOS cameras. The processing of this kind of data is a challenging task for these resource-constrained units (i.e., energy) which are used in practice to deal with only scalar data. In fact, sensor nodes are typically battery powered; and changing this component is in general undesirable and it is even impossible in some applications. In conventional wireless sensor networks the processing's energy consumption is neglected since this operation is considered as simple. However, in a WVSN the captured video stream is compressed before transmission; according to the literature [1], the video compression can consume about 2/3 of the total power for video communications over wireless channels. Therefore, energy efficient compression is required while maintaining acceptable video quality.

In [2], we propose an energy efficient adaptive video compression scheme for WVSNs. This solution, which is based on the H.264/AVC standard with only intra prediction, operates using two modes, namely the standby mode and the rush mode. In the standby mode, a low frame rate (FR) is adopted while the image quality is maintained to a high level. In the rush mode, the FR is increased and the Flexible Macroblock Ordering (FMO) [3] coding tool provided by the H.264/AVC video compression standard is used to produce two service-differentiated macroblocks categories: region of interest (ROI) and background (BKGD). In addition, to decrease the transmission's energy consumption, a controlled bit rate adaptation, based on frequency selectivity is applied to reduce the background data amount. Finally, CAVLC entropy encoding is used. In fact, even if it is slightly less efficient than CABAC, it can approach the entropy of the source with a reduced complexity. The intra mode is used because it is adapted to the WVSN context. In fact,the authors in [4] propose a comparative study concerning the H.264/AVC intra mode against JPEG and JPEG2000. They conclude that the H.264/AVC intra mode offers an interesting compromise in terms of complexity, quality and coding efficiency. In addition, as will be shown in this paper, the inter-image coding mode consumes much more energy. In fact, considerable energy can be saved when the H.264/AVC intra-only mode (reducing the encoding energy [5]), and energy efficient transmission schemes are adopted. Consequently, this makes the H.264/AVC intra-only mode an attractive tool for video compression in a WVSN. We demonstrate in [2] that the proposed solution is able to extend network lifetime up to three times compared to the classical scheme, thanks to the important amount of energy that is saved during the transmission phase. However, the consumed energy during the compression should also be considered in the WVSN context. Therefore, a mathematical model is needed in order to predict the consumed energy by our encoder. In [6], the authors propose a power-rate-distortion (P-R-D) model which is widely used to estimate the consumed energy during the video compression [7, 8]. However, in [9], the authors high-

978-1-4799-5751-4/14/$31.00 ©2014 IEEE

light that their analytical approach described in [6] cannot be easily extended to other video encoders such as H.264/AVC standard. Hence, they propose an operational approach for off-line energy consumption analysis and modelling which can be applied to generic video encoders. It relies on the consideration of the parameters that are responsible of consuming more or less energy during the video signal compression. In our case, the amplitude resolution that is controlled by the quantization parameter (QP) and the temporal resolution represented by the FR are considered both.
Therefore, in this paper we develop an energy consumption model for an H.264/AVC intra-only video encoder considering the QP and FR parameters. Intuitively, one can say that the energy decreases when lowering the FR. The same behaviour is expected when the QP is increased. However, the important question that we are trying to answer in this paper is: How does the energy decrease when only intra-image coding is used? The answer will help to model the consumed energy by any H.264/AVC intra-only mode based encoder, and will further allow us to optimize the consumed energy in WVSNs [2].

2. PROPOSED MODEL

In order to model the consumed energy during the video signal compression, extensive tests, using four different CIF and QCIF video sequences encoded at 14 different QPs and 10 different FRs, have been conducted using JM18.4 implementation of the H.264/AVC video coding standard in its baseline profile. For accurate measures we do not use the frequency selectivity based bit rate adaptation neither the FMO [2]. Hence, we use the classical JM18.4 implementation. In addition, we control the running of the encoder in order to be in a real-time fashion and make it the only executed process in one microprocessor.
The consumed energy for processing can be modelled as presented in [10] as a function of the the number of clock cycles by:

$$E_{Proc}(N) = N_{cyc} C_{total} V_{dd}^2 + V_{dd}(I_0 e^{\frac{V_{dd}}{nV_T}})(\frac{N_{cyc}}{f}) \quad (1)$$

Where N_{cyc} is the number of clock cycles, C_{total} is the average capacitance switched per cycle, V_{dd} is the supply voltage, I_0 is the leakage current, f is the clock speed, V_T is the thermal voltage and n a processor dependent constant. On the other hand, since:

$$T_{Proc} = \frac{N_{cyc}}{f} \quad (2)$$

then,

$$N_{cyc} = T_{Proc} f \quad (3)$$

Where T_{Proc} is the processing time, which is the encoding time in our case returned by the encoder.

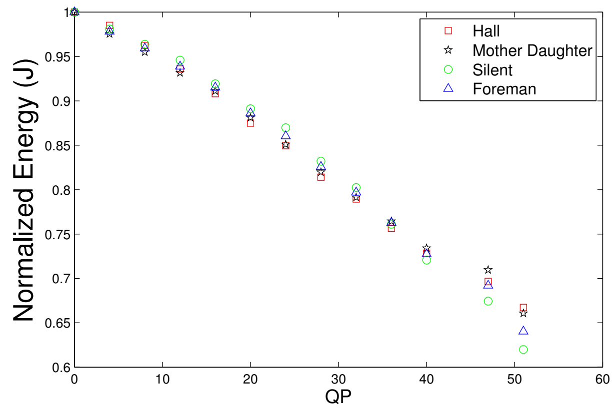


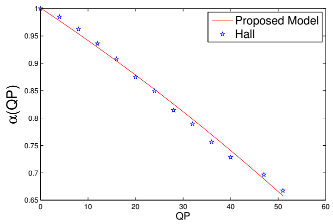
Fig. 1. Normalized consumed energy for QCIF tested sequences

2.1. Modelling the consumed energy as a function of QP

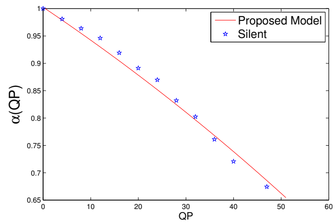
Fig. 1 illustrates the behaviour of the normalized consumed energy during the video compression of the tested sequences. We can notice that the energy decreases slowly when the QP is increased. This can be explained by when increasing the QP value, the quantization becomes more severe, generating macroblocks with more null coefficients. Consequently, the energy consumed during the CAVLC entropy coding is decreased. The only difference between the CIF and QCIF sequences is the amount of the consumed energy, but the shape is the same for both resolutions. Actually, Fig. 1 shows the behaviour of a reduction factor that is QP-dependent. We name it $\alpha(QP)$. This factor reaches its maximal value of 1 at $QP = QP_{min} = 0$, then slowly decreases while the QP increases. Based on the above mentioned arguments, we propose to model $\alpha(.)$ as follows:

$$\alpha(QP) = 2 - exp(a \times QP) \quad (4)$$

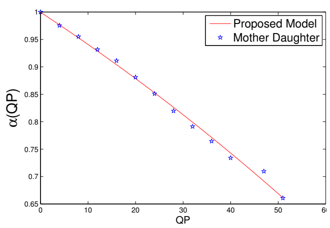
Where a is a content-dependent coefficient obtained by minimizing the root mean squared error (RMSE) between measured and predicted data. As shown in Fig. 2, the proposed model in Eq. 4 can predict the reduction factor $\alpha(QP)$ accurately. We were curious about the behaviour of $\alpha(.)$ using the inter-prediction mode (i.e GOP IPPPP), and hence take advantage of the temporal correlation that characterizes the video signal. Fig. 3 presents three results. The first is that the consumed energy by the compression decreases differently from one mode to another (compared with Fig. 1). The second is that there is a significant difference from a sequence to another (Fig. 3 (b)). This is due to the inherent quantity of motion in each sequence. The third is that using inter-image prediction consumes much more energy (Fig. 3 (a)). Similar results are reported in [5] where the authors ran simulations on Stargate nodes.



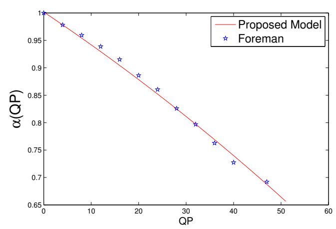
(a) Hall: a=0.005722



(b) Silent: a=0.005839

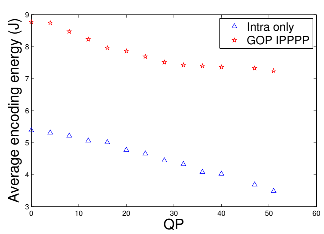


(c) Mother Daughter: a=0.005797

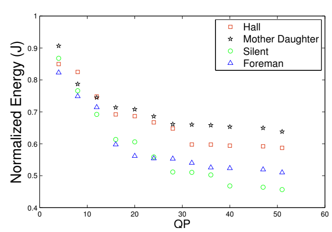


(d) Foreman: a=0.005831

Fig. 2. Measured data and its approximation using the proposed model in Eq. 4



(a) Average encoding energy



(b) Normalized energy

Fig. 3. Normalized consumed energy for four inter-coded sequences

2.2. Modelling the consumed energy as a function of FR

The FR is the second factor that we consider to predict the energy consumption of our video encoder. We can change the FR in the JM implementation of H.264/AVC by varying of the parameter named Frame Skip (FS). The relation between these two parameters is given as follows:

$$FS = \lceil \frac{FR_{max}}{FR} \rceil - 1 \qquad (5)$$

Where FR_{max} is the maximal FR allowed, and $\lceil . \rceil$ is the ceiling function.

Fig. 4 presents the measured normalised energy while varying the FS. As can be seen, the energy consumed decreases when increasing the FS and thus decreasing the FR. This can obviously be explained by when reducing the FR, the FS is increased and hence less frames are encoded leading to a considerable reduction of the consumed energy. Note that the FR_{max} of the tested video sequences is 30 fps and using Eq.

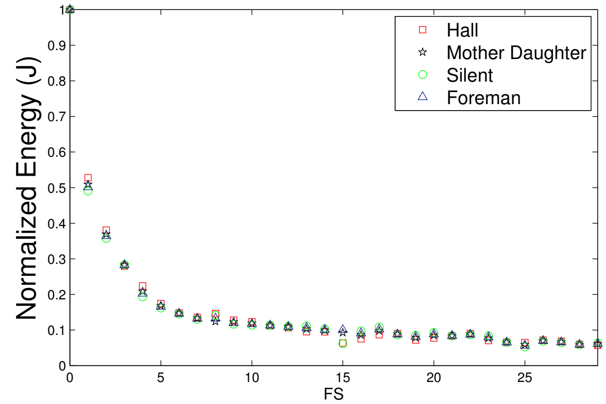


Fig. 4. Normalized consumed energy for QCIF tested sequences

5 one can deduce the tested FRs. Also, we notice that the incrementation of the FS by one reduces the energy by about the half and so on, which is an obvious and expectable behaviour. Similar to the previous subsection, we obtain the same behaviour for CIF resolution but the amount of the consumed energy is larger. In fact, Fig. 4 shows the behaviour of a reduction factor that is FS-dependent. We name it $\beta(FS)$. This factor actually reaches its maximal value of 1 at $FS = FS_{min} = 0$ and quickly decreases toward its minimal value at $FS = FS_{max} = 29$. In addition, this factor does not attain the zero value since there is at least one frame to be encoded. Based on the above mentioned arguments we propose to model $\beta(.)$ as follows:

$$\beta(FS) = \frac{1}{2^{FS}} + b \qquad (6)$$

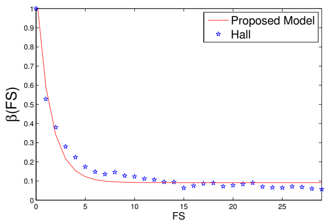
Where b is a content dependent parameter obtained by minimizing the root mean squared error (RMSE) between measured and predicted data. As shown in Fig. 5, the proposed model in Eq. 6 can model the reduction factor $\beta(FS)$ accurately.

2.3. The global model

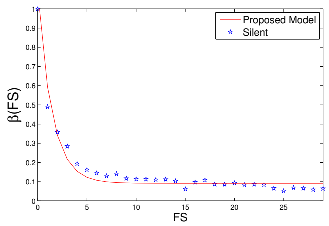
The maximal energy consumed during the video signal encoding is reduced by a QP-dependent reduction coefficient then by a FR-dependent reduction coefficient. Hence, the global model of the consumed energy during the video signal encoding is given by:

$$E(QP, FR) = E_{max} \times \alpha(QP) \times \beta(\lceil \frac{FR_{max}}{FR} \rceil - 1) \qquad (7)$$

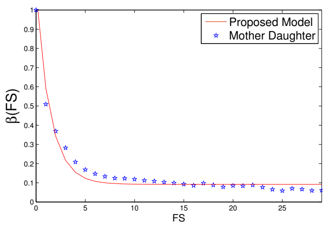
Where E_{max} is the maximal energy consumed during the video signal compression that is reached for the couple (QP_{min}, FR_{max}). E_{max} depends also of the microprocessor that runs the encoder. This model can help to predict



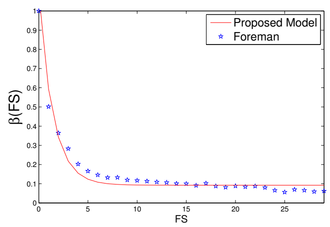
(a) Hall: b=0.0924



(b) Silent: b=0.0915



(c) Mother Daughter: b=0.0921



(d) Foreman: b=0.0922

Fig. 5. Measured data and its approximation using the proposed model in Eq. 6

the energy consumption of any encoder based on H.264/AVC using intra-prediction only.

3. MODEL APPLICATION

In this section we apply the proposed model to our encoder proposed in [2]. The parameters a and b are set to their average values respectively 0.0058 and 0.0921. We use a machine with an Intel 2.93 GHz Core 2 Duo processor. The different coefficients of Eq. 1 are available on Intel's website [11]. The measured consumed energy is obtained by first injecting the encoding time returned by the encoder in Eq. 3. Then, the result is inserted in Eq. 1. The different tested configurations are reported in Table 1. In this table, FR_S and FR_R refer to the considered FR during the standby mode and the rush mode respectively. As explained in [2], we assume that the WVSN using the proposed encoder includes an intelligent motion detection system. Actually, this system is responsible of commanding the video nodes to switch from one mode to another. Therefore, we report in Table 1, the percentage of the video sequence that is considered in each mode. For example, the first 30% of the sequence Bus is considered as standby mode, while the remaining 70% is considered as rush mode. The prediction error (PE) of the proposed model for the tested video sequences under the different configurations is less than 6%, as stated in Table 1.
For more precision, we present in Fig. 6 detailed information about the measurements of the consumed energy by our encoder and the predicted energy using the proposed model in Eq. 7 for each mode. As can be seen, the proposed model predicts the consumed energy accurately.

Table 1. Different tested configurations and the PE

Seq.	Res.	FR_{max}	QP	FR_S	FR_R	PE
Bus	CIF	30	36	3.75 (30%)	10	5%
Soccer	QCIF	15	4	5 (60%)	15	4%
City	CIF	30	30	5 (35%)	10	2%
Mobile	QCIF	50	10	10 (50%)	15	4%

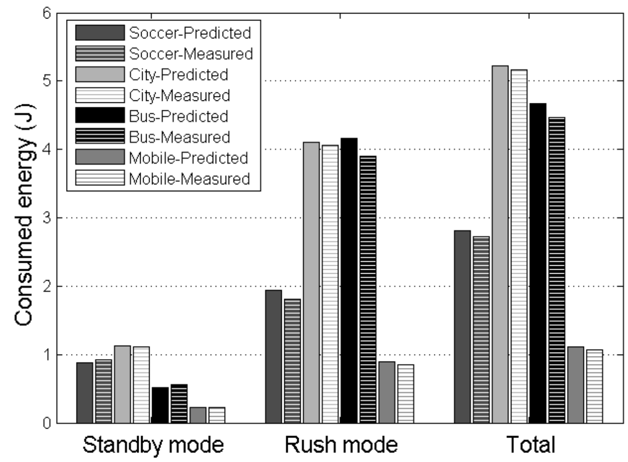


Fig. 6. The measured and the predicted energy for each mode and all the sequences

4. CONCLUSION

In this paper, an energy consumption analysis and modelling of H.264/AVC intra-only mode based encoder dedicated to WVSNs is presented. Throughout our analysis, we show that the consumed energy during the compression is in fact the maximal energy that is reached for the couple (QP_{min}, FR_{max}), which is reduced by a QP-dependent coefficient, then by a FR-dependent coefficient. In addition, we compare the energy consumption of the intra-only mode against the inter mode (i.e. GOP IPPPP) and prove the energy efficiency of the first mode. Hence, combining this mode to energy efficient transmission schemes can save considerable energy in a video sensor node. Finally, the proposed model is applied on our encoder under different configurations (i.e. resolution, QP and FR). This model provides accurate predictions with an average error of 4%.

5. REFERENCES

[1] X. Lu, Y. Wang, and E. Erkip, "Power efficient h. 263 video transmission over wireless channels," in *Proceedings of International Conference on Image Processing (ICIP)*, Rochester, NY, USA, September 2002, vol. 1, pp. 533–536.

[2] O. Alaoui-Fdili, F.X. Coudoux, Y. Fakhri, P. Corlay, and D. Aboutajdine, "Energy efficient adaptive video compression scheme for wvsns," in *Proceedings of European Signal Processing Conference (EUSIPCO)*, Marrakech, Morocco, September 2013.

[3] S. Wenger and M. Horowitz, "FMO: Flexible Macroblock Ordering," *JVT-C089*, 2002.

[4] A. Al, B.P. Rao, S.S. Kudva, S. Babu, D. Sumam, and A.V. Rao, "Quality and complexity comparison of h. 264 intra mode with jpeg2000 and jpeg," in *Proceeding of International Conference on Image Processing (ICIP)*, Singapore, October 2004, vol. 1, pp. 525–528.

[5] S. Ullah, J.J. Ahmad, J. Khalid, and S.A. Khayam, "Energy and distortion analysis of video compression schemes for wireless video sensor networks," in *Proceedings of Military Communications Conference (MILCOM)*, Baltimore, MD, USA, November 2011, pp. 822–82.

[6] Z. He, Y. Liang, L. Chen, I. Ahmad, and D. Wu, "Power-rate-distortion analysis for wireless video communication under energy constraints," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 15, no. 5, pp. 645–658, 2005.

[7] Y. He, I. Lee, and L. Guan, "Distributed algorithms for network lifetime maximization in wireless visual sensor networks," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 19, no. 5, pp. 704–718, 2009.

[8] J. Zou, H. Xiong, C. Li, R. Zhang, and Z. He, "Lifetime and distortion optimization with joint source/channel rate adaptation and network coding-based error control in wireless video sensor networks," *IEEE Trans. on Vehicular Technology*, vol. 60, no. 3, pp. 1182–1194, 2011.

[9] Z. He, J. Eggert, W. Cheng, X. Zhao, J. Millspaugh, R. Moll, J. Beringer, and J. Sartwell, "Energy-aware portable video communication system design for wildlife activity monitoring," *IEEE Circuits and Systems Magazine*, vol. 8, no. 2, pp. 25–37, 2008.

[10] D. Rui, W. Pu, and I.F. Akyildiz, "Correlation-aware qos routing with differential coding for wireless video sensor networks," *IEEE Trans. on Multimedia*, vol. 14, no. 5, pp. 1469–1479, 2012.

[11] Intel Core2 Duo Processor, "`http://support.intel.co.jp/pressroom/kits/core2duo/`," [Online; accessed 07-February-2014].