DECODER ENERGY-AWARE INTRA-CODED HEVC BIT STREAM GENERATION

Thanuja Mallikarachchi, Dumidu S. Talagala, Hemantha Kodikara Arachchi, Anil Fernando

Centre for Vision Speech and Signal Processing, University of Surrey, United Kingdom.
{d.mallikarachchi, d.talagala, h.kodikaraarachchi, w.fernando}@surrey.ac.uk

ABSTRACT

The rising demand for media consumption via mobile devices and the emergence of complex video coding algorithms present an additional challenge for the energy management algorithms in resource-constrained consumer electronic devices. Thus, generating energy-optimized video bit streams will positively contribute towards overcoming this challenge. This paper introduces a novel energy model for intra-frame decoding of HEVC encoded video that predicts the decoding energy of a coding unit. Thereafter, a novel energy-rate-distortion optimized coding mode selection algorithm is proposed to generate energy-efficient bit streams using the proposed energy model within the encoder. The proposed energy model is shown to predict the decoding energy of a coding unit with an average error less than 2%. Moreover, the proposed coding mode selection algorithm achieves an average 10.8% reduction in the energy consumed at the decoder with a -0.25 dB impact to the Bjøntegaard Delta-Peak Signal-to-Noise Ratio.

Index Terms— HEVC intra-prediction, decoder complexity, decoder energy optimization, energy-efficient bit streams.

1. INTRODUCTION

Multimedia consumption is by far one of the most resource intensive operations performed by handheld Consumer Electronic (CE) devices (e.g., smart phones, tablets, etc.) due to the limited availability of processing and energy resources. The recent advancements in media formats and technologies, together with the growth in mobile video data (expected to reach three-fourths of the worlds overall mobile data traffic by 2019 [1]), have contributed towards the disproportionately large utilization of a device's resources. Thus, the media consumption-driven energy use of these devices has become a critical bottleneck that must be overcome.

Energy consumed by multimedia processing, i.e., video encoding, decoding and presentation, is tightly coupled to the complexity of the codec as well as the content. For example, the recent proliferation of High Definition (HD) and Ultra High Definition (UHD) video content demonstrably increases the processing required, and coincidentally demands more complex video compression algorithms, both of which adversely affect a device's energy consumption. Yet, although the processing capabilities of devices have kept up with Moore's law, battery capacity increases have lagged far behind. The increased complexity of novel coding standards such as High Efficiency Video Coding (HEVC) [2] [3] therefore profoundly impact on the energy consumption of CE devices, especially in the case of high resolution content. In this context, a decoder energy-aware video encoding framework could significantly reduce the energy consumption of consumer devices, thereby maximizing the consumers' overall quality of experience of both the devices and the content in future multimedia applications.

Initiatives such as Green-MPEG [4] and technologies such as scalable and adaptive streaming have spawned numerous research outputs on energy-efficient video stream generation for resource constrained devices. However, the state-of-the-art solutions do not consider the intricacies associated with the novel coding modes and features in HEVC; thus, the scope and potential exists to investigate how to generate energy-efficient bit streams (within the video coding layer) that minimally impact the coding efficiency. To this end, this paper proposes a framework to model the energy utilization of decoding intra-coded frames in HEVC. The model identifies the key energy-intensive operations and models the energy consumption, thereby providing a mechanism to predict the energy required to decode a Coding Unit (CU). An energy-rate-distortion optimized mode selection mechanism, which leverages the proposed energy consumption model, is introduced thereafter to determine the optimum coding mode configuration that minimizes the decoding energy requirements of a CU with minimal impact on the coding efficiency.

The remainder of the paper is organized as follows. Sec. 2 presents an overview of the related work, while Sec. 3 describes the proposed energy prediction model for decoding of HEVC intra-frames. The decoder energy-optimized coding mode selection is described in Sec. 4, and is followed by a discussion of the experimental results in Sec. 5. Finally, Sec. 6 concludes with a summary and the implications of this study.

2. BACKGROUND AND RELATED WORK

The survey presented by Hoque *et al.* [5] groups energy-efficient media streaming solutions into two categories; solutions that optimize the wireless receiving energy with unmodified content, and those that modify the content in order to reduce the total energy consumed by the wireless receiver and the decoder. The focus of this work revolves around the codec and the application layer; hence, the following section mainly describes the background and related work that pertains to the second category.

The generation of energy-optimized bit streams at the encoder requires an accurate and detailed model that predicts the energy consumed by the decoder. An energy use assessment of the HEVC encoder, based on the assembly level instruction analysis, is presented by Saab *et al.* [6], and Henkel *et al.* [7] identify energy consuming nodes within the HEVC encoding loop. However, these works focus only on the encoder thus, the results cannot therefore be directly applied to a decoder energy-aware encoding process. The level of detail in the energy model proposed by Herglotz *et al.* [8] is shown to be insufficient to perform a complete energy-rate-distortion optimization at the CU level, which is also the drawback associated with

This work was supported by the ACTION-TV project, which is funded under European Commissions 7^{th} Framework Program (Grant number: 611761).

the work presented in [9], that studies the relationship of the decoder complexity, the content and the Quantization Parameter (QP). In terms of the energy reduction in intra-predicted frames, Nogues *et al.* [10] proposed a modified decoder that skips the in-loop filtering process based on a desired activation level. However, the method does not consider the diversity of the energy consumed by the HEVC coding tools, and in addition, the modifications required to the decoder may hinder the usability of the method with the established hardware implementations.

On the power management front, technologies such as Dynamic Voltage and Frequency Scaling (DVFS) and Dynamic Power Management (DPM) have spawned much of the research output on efficient power utilization. For example, the method proposed by Guo *et al.* [11] models the Central Processing Unit (CPU) power consumption, and Liang *et al.* [12] make use of hardware performance counters, to predict the decoding complexity and apply DVFS techniques afterwards to reduce the power consumption. However, in general, these methods suffer from increased frame drops that affect the users' quality of experience, especially in the case of high frame rate content [12]. In the literature, energy-aware video encoding and delivery mechanisms typically adopt scalable video coding architectures that involve proxy servers such as MANEs, media transcoding solutions [13], and dynamic adaptive streaming technologies such as MPEG-DASH [14]. In general, these solutions utilize enhancement layers with incremental energy levels, dynamic changes to the coding structure and multiple streams with different energy levels to manage the energy consumption on-demand. However, they are limited to parameters such as the QP, spatial resolution, and frame rate, and therefore novel features in the HEVC architecture have not been considered as potential parameters for exploitation to alter the energy required to decode a stream.

3. ENERGY UTILIZATION MODELING

Energy profiling of the encoding and decoding operations in HEVC have been attempted using both direct measurements of the voltages and currents [8] as well as using instruction level profiling methods [6, 15]. The latter, which is capable of providing CPU instruction count estimates for each function, has been used extensively for the purposes of profiling and complexity analysis [6]. The proposed model therefore utilizes a similar profiling mechanism for the analysis of the decoder operations. Here, the estimated CPU cycles that are obtained for each operation is mapped to the energy consumed by a particular operation using the relationship between the clock frequency, the effective switched capacitance C_{EFF} and the supply voltage V [16]. Thus, the energy consumption of the decoder, denoted by E, can be expressed as

$$E = C_{EFF} \times V^2 \times C, \tag{1}$$

where C is the estimated CPU cycles. The energy consumed therefore exhibits a linear relation to the number of CPU cycles. Hence, in this context, the number of CPU execution cycles consumed by the decoder's operations can be considered as a substitute for the energy it consumes, and is used throughout the course of this work.

3.1. Formulating the energy model

The process of reconstructing a CU at the decoder consists of two phases; the decoding phase and the decompression phase. The decoding phase, which performs the entropy decoding of the syntax and residual coefficients, is composed of five main steps; the prediction mode, PU size, *luma* mode, *chroma* mode and coefficient decoding, of which the coefficient decoding phase consumes a greater portion of the CPU time. The decompression phase, which performs the predicting and reconstructing based on the decoded information, is composed of operations such as a filtering check, reference sample handling, prediction and inverse transform processes.

A CU's decoding energy E_{dec} can therefore be expressed as

$$E_{dec} = e_{pmode} + e_{psize} + e_{lmode} + e_{cmode} + e_{coeff}, \tag{2}$$

where e_{pmode}, e_{psize}, e_{lmode}, e_{cmode}, and e_{coeff} are the complexity estimates for decoding the prediction mode, PU size, *luma* mode, *chroma* mode, and the transform coefficient decoding, respectively. The decompression energy E_{dcomp} can be expressed similarly as

$$E_{dcomp} = \sum_{i=1}^{N} \Big\{ e_{rf}(i) + e_{lm}(i) + 2 \times e_{ch}(i) + e_{df}(i) + e_{it}(i) \Big\}, \tag{3}$$

where e_{rf}, e_{lm}, e_{ch}, e_{df}, and e_{it} correspond to the estimated complexities for handling of the reference samples, *luma* and *chroma* predictions, DC filtering, and the inverse transform operations, respectively. In this case, e_{it} is considered only for the Transform Units (TU) for which the Coded Block Flag (CBF) is non-zero and N is the total number of TUs in the CU. It should however be noted that computing energy consumed becomes challenging when the TU structure within the CU becomes complex, e.g., when a TU tree exists within the CU. In such cases, E_{dcomp} should be calculated for each TU separately.

Thus, the energy required to process the j^{th} CU is given by

$$E_{CU}(j) = E_{dec} + E_{dcomp} + \sum_{k=1}^{K} e_{overhead}(k) \times \frac{nTU(k)}{nTTU(k)}. \tag{4}$$

In this context, K denotes the number of unique TU sizes within the j^{th} CU, $nTU(k)$ is the number of TUs that correspond to the k^{th} TU size, $nTTU(k)$ is the total number of TUs (of similar size to the k^{th} TU size) that could occur within a CTU, and $e_{overhead}(k)$ encompasses the overhead due to the quadtree structure of the k^{th} TU size. For example, if 8×8 TUs exist in the j^{th} CU, $nTTU(k)$ is 64 and $nTU(k)$ is the number of 8×8 TUs that actually exist within the j^{th} CU. Then the decoding energy for the CTU, E_{CTU} is given by the summation of $E_{CU}(j)$ values, and can be expressed as

$$E_{CTU} = \sum_{j=1}^{M} E_{CU}(j), \tag{5}$$

where M is the total number of CUs that constitute the CTU.

3.2. Decoder profiling for intra-predicted HEVC video frames

For analysis purposes, video sequences that exhibit diverse spatial characteristics have been selected as training and validation sequences[1]. Here, the decoding complexity of an intra-coded CU is analyzed, and is presented in terms of the number of consumed CPU cycles (with respect to an Intel core i7-2600 CPU with 3.4 GHz, 8GB RAM reference system using a x86 architecture). The numerical results presented in the following subsections therefore specify the estimated CPU cycles of the HM 16.0 decoder [17] running on

[1]The sequences used in this analysis can be found in Table 2.

Table 1. Estimated CPU cycles of the decompression operations in the decoder.

TU size	Prediction Process (e_{lm})									e_{rf}		e_{it}
	DC	Planar	Vertical	Horizontal	Integer Angles			Fractions				
	1	0	26	10	2	18	34	Hor	ver	Filt.	Non-Filt.	
32	13707	44097	24332	42211	40612	24332	22733	47510	29632	13496	12152	694982
16	3835	11857	8294	12892	11192	8293	6593	12971	8376	8021	7284	66797
8	1203	3417	2894	4166	3392	2893	2169	3858	2636	4750	4360	15306
4	463	1117	1201	1552	1552	1201	869	1346	995	-	3418	8518

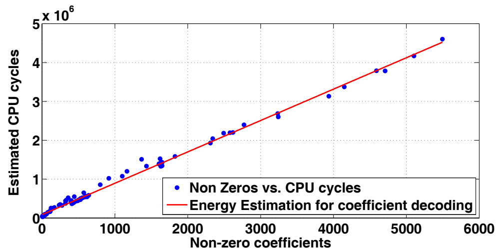


Fig. 1. Modeling of estimated CPU cycles with the number of non-zero coefficients identified within a CTU.

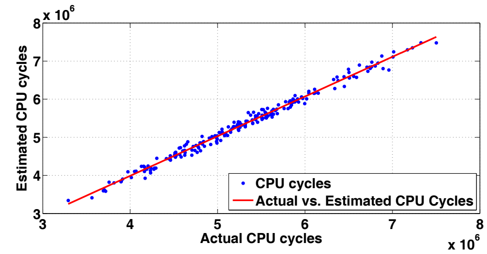


Fig. 2. Actual and predicted CPU cycles per CTU for the Cafe 1080p sequence for QPs 22, 27, 32, 37.

the reference system, and represents the relative complexity level of each operation, and by extension, its energy consumption.

The experimental results reveal that the CPU cycles consumed by the decoding operations, other than the transform coefficient decoding process denoted by e_{coeff}, are comparatively small (typically $< 1\%$). Therefore, the proposed energy model is simplified to utilize the CPU cycles corresponding to e_{coeff}, as a representation of the decoding energy E_{dec} of the CU. Furthermore, the empirical analysis suggests that the transform coefficient decoding energy can be modeled as a function of the number of non-zero transform coefficients; a measurement that varies with the QP and the content. Thus, the data obtained from the training set can be fitted to a linear model, shown in Fig. 1, given by

$$e_{coeff} = 818.2 \times x + (1.039 \times 10^5), \quad (6)$$

where x is the number of non-zero transform coefficients. The fitted curve exhibits a R-Square Goodness-of-Fit of 0.995 with respect to the validation set, which suggests that a general behavior is being modeled.

The decoder complexity relevant to the decompression operations predominantly depend on the TU sizes and the prediction modes selected for the CU. For example, the complexity of reference sample handling denoted by e_{rf} (i.e., reference sample filtering and the filling of the reference sample arrays) depends on the TU size, availability of the neighboring samples and the result of the filter check process. Furthermore, the intra-prediction architecture in HEVC defines the TU sizes and the corresponding prediction modes that require filtering during the decoding process [18]. Hence, the complexity of the decoding process that pertains to e_{rf}, when each TU block is either filtered or not-filtered depending on the selected mode [18], can be summarized as shown in Table 1.

Similarly, the energy consumption for the prediction operations vary based on a number of factors such as the interpolation requirements, TU size, prediction direction, etc. [18], as seen in Table 1[2]. In addition, the DC prediction in *luma* TUs smaller than 32 $\times$ 32 requires filtering [18], for which the CPU cycles are estimated to be

[2]The number of CPU cycles for the *chroma* decompression is deduced by association with the TU size of the *chroma* channel, i.e., the *chroma* TU width is half of the *luma* TU width for the 4:2:0 chroma sampling format.

Table 2. Energy prediction performance of the proposed model.

Training Set		Validation Set	
Sequence	$P_e(\%)$	Sequence	$P_e(\%)$
Akiyo	1.54	Coastguard	0.59
Bridge-far	1.47	Container	0.56
Waterfall	0.88	BasketballDrill	0.77
BasketballPass	0.61	Poznan St. HD	0.44
Band HD	1.85	Beergarden HD	0.32
Kimono HD	0.67	Dancer HD	0.44
GT Fly	0.94	Cafe HD	0.84
		Musicians HD	0.10
		Traffic UHD	1.01
		MenPlants UHD	1.21
Average	**1.13**	**Average**	**0.58**

1075, 539 and 271 for the 16 $\times$ 16, 8 $\times$ 8 and 4 $\times$ 4 TU sizes, respectively. The complexity of the inverse transform, denoted by e_{it}, that occurs when the Transform Skip Mode (TSM) is inactive [2] is also presented in Table 1.

Finally, the total CPU cycles consumed by these main decoder operations need to be supplemented with an additional offset for the overhead caused by the quadtree structure of splitting CUs and TUs into the multiple depth levels. The profiling reveals that this overhead can be presented as a function of TU size; hence, CPU cycle offsets of 720385, 907029, 1706934, and 3133889 can be mapped to the TU sizes 32, 16, 8 and 4, respectively.

3.3. Verification of the energy model

Verification of the proposed model is performed using a set of validation sequences which are independent from the training sequences. The sequences have been encoded using the HM 16.0 reference software [17] in the *All Intra Main* configuration with QPs 22, 27, 32, and 37 and the energy consumption for the decoding operations are estimated in terms of the CPU cycles [15]. The cumulative number of CPU cycles are thereafter compared with the CPU cycles predicted from the proposed model using the average prediction error given by

$$P_e = 100 \times \frac{|E_a - E_p|}{E_a}, \quad (7)$$

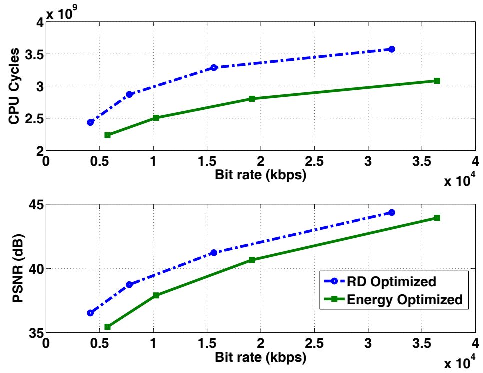


Fig. 3. The RD performance and energy saving graphs for the RD optimized bit stream and the energy optimized bit stream.

where E_a and E_p are the actual CPU cycles [15] and the predicted CPU cycles from the proposed model, respectively. The verification results for the proposed model are presented in the Fig. 2, and should ideally exhibit a linear relationship. A summary of the average P_e for all the sequences are presented in Table 2. The average error observed is less than 2%, which suggests that the prediction is more accurate compared to the 3.2% error reported by Herglotz *et al.*[8]. The proposed model achieving a smaller P_e lays the foundation for the energy-optimized HEVC intra-coding framework described in the following section.

4. DECODER ENERGY OPTIMIZED INTRA-CODING IN HEVC

Next, a mechanism to utilize the proposed energy model to select energy-rate-distortion optimized coding modes during encoding is presented.

4.1. Energy-rate-distortion optimization

The Rate-Distortion (RD) optimization performed by the encoder determines the selection of the coding modes and the structure to be applied to the content. The Lagrangian cost function evaluated therein can be expressed as

$$\min \; D(p) + \lambda R(p) \; \big| \; p \in P_k, \tag{8}$$

where $\lambda \geq 0$ is the Lagrange multiplier, p is a particular coding parameter combination in the set of all the possible coding options P_k, and $D(p)$, $R(p)$ are the distortion and rate associated with the selected set of coding parameters, respectively. The minimum cost coding mode and structure, for a given CU, contributes towards achieving the best coding efficiency within the encoder. However, although optimized in a rate-distortion sense, it may still consume significant energy to decode, whereas, another mode and a structure may consume less energy albeit at a higher RD cost. The RD and energy consumption graphs for a single frame of Poznan Street 1088p sequence, presented in the Fig. 3, illustrate this trade-off. In this case, the energy efficient stream is obtained by simply utilizing the decoding energy as the only parameter in the Lagrangian cost function given by

$$\min \; E(p) \; \big| \; p \in P_k, \tag{9}$$

where, $E(p)$ represents the cost of decoder energy for a particular coding parameter combination. The loss in coding efficiency between the RD-optimized and the energy-optimized streams, in terms of the Bjøntegaard-Delta PSNR (BD-PSNR) [19], is -1.618 dB, while the average energy saving achieved with the latter is 12.3%.

This leads to the conclusion that mere energy optimization would result in a bit stream which is RD inefficient. Therefore, the consideration of all three parameters, i.e., decoding energy, rate and distortion, is essential during the mode and coding structure selection. Thus, the modified Lagrangian cost function can be expressed as

$$\min \; D(p) + \lambda R(p) + \lambda_e E(p) \; \big| \; p \in P_k, \tag{10}$$

where $\lambda_e \geq 0$, determines the trade-off between balancing the impact of energy efficiency with the RD efficiency.

4.2. Determining the energy trade-off factor λ_e

Evaluating the impact of a variety of λ_e, it can be observed that it affects both the energy saving and the RD performance. The experiments conducted using the training sequences in the Table 2, with λ_e ranging from 0 to 1, show that the energy saving that can be achieved increases with increasing λ_e. This however increases the bit rate, significantly impacting the RD performance, and the energy consumption of the radio transceiver managing the downlink data reception [5] of the CE device running the decoder.

Previous research in the literature suggests that the power consumption of the radio receiver scales linearly with respect to the data rate. Thus, the energy consumed to receive a CU for a sequence encoded at a frame rate F can be defined as,

$$E_{rec}(p) = \{\alpha \times R(p) + \beta\} \times \frac{1}{F}, \tag{11}$$

where $E_{rec}(p)$ is the energy consumption for data reception, $R(p)$ is the data rate associated with the p^{th} coding mode, and α and β are constants. Applying a linear model [20] with respect to LTE systems, α and β were determined to be 0.003 and 1.98×10^{-3}, respectively. Hence, the absolute energy saving that can be achieved with the energy-rate-distortion optimization is given by

$$\Delta E_{abs} = \{E(p) + E_{rec}(p)\} - \{E(p_o) + E_{rec}(p_o)\}, \tag{12}$$

where p and p_o are the set of coding modes selected by the encoder when using (8) and (10) as the optimization cost function for mode selection, respectively.

In order to empirically determine λ_e that maximizes (12), the test sequences were encoded using the HM reference software using the *All Intra Main* configuration for 50 frames and QPs 22, 27, 32, and 37 using the energy-rate-distortion optimization in (10). Fig. 4 illustrates the resulting averaged variation of ΔE_{abs} with respect to the corresponding λ_e values. The observed behavior suggests that the λ_e values that range from 0.001 to 0.005 maximize ΔE_{abs}, which considers the energy required for both the decoding and the data reception. An averaged value λ_e of 0.0028 is therefore selected to maximize ΔE_{abs} throughout the course of this work.

5. EXPERIMENTAL RESULTS AND DISCUSSION

The performance of the proposed coding mode selection algorithm is first analyzed utilizing the energy model described in Sec. 3. A similar analysis is performed thereafter, by utlizing the energy model proposed by Herglotz *et al.* [8]. Finally, the performance of the proposed method, when supplemented by a dynamic in-loop filter manipulation algorithm is also discussed to illustrate its extensibility to further reduce the decoder energy consumption. The performance is analyzed and presented for the test sequences encoded using the *All*

Table 3. RD and energy saving performance of the proposed method.

	Sequence	Proposed		Proposed + in-loop filtering		Herglotz *et al.*[8]	
		$\Delta E_{abs}(\%)$	BD-PSNR (dB)	$\Delta E_{abs}(\%)$	BD-PSNR (dB)	$\Delta E_{abs}(\%)$	BD-PSNR (dB)
Training Set	Akiyo	11.97	-0.35	26.02	-0.40	-2.51	-0.18
	Bridge-far	12.70	-0.15	22.50	-0.15	-0.08	-0.08
	Waterfall	9.53	-0.15	16.48	-0.15	1.49	-0.10
	BasketballPass	10.67	-0.19	23.05	-0.19	1.24	-0.26
	Band HD	9.53	-0.26	21.26	-0.26	-4.72	-0.14
	Kimono HD	6.08	-0.24	16.63	-0.26	-7.00	-0.22
	GT Fly	14.11	-0.23	28.01	-0.24	-3.19	-0.23
	Average	**10.65**	**-0.22**	**21.99**	**-0.23**	**-2.11**	**-0.17**
Validation Set	Coastguard	9.73	-0.18	17.06	-0.18	2.94	-0.04
	Container	10.73	-0.18	22.04	-0.18	1.94	-0.19
	BasketballDrill	12.28	-0.29	23.99	-0.29	3.49	-0.29
	Poznan St.	11.84	-0.25	21.43	-0.25	3.95	-0.22
	Beergarden	9.37	-0.37	20.73	-0.37	-4.59	-0.25
	Dancer	13.88	-0.23	24.52	-0.32	6.68	-0.14
	Cafe	8.01	-0.29	20.20	-0.40	-1.65	-0.31
	Musicians	10.07	-0.29	22.38	-0.29	-3.02	-0.24
	Traffic	13.26	-0.34	25.27	-0.30	3.36	-0.29
	MenPlants	10.41	-0.45	21.48	-0.45	-3.32	-0.25
	Average	**10.96**	**-0.28**	**21.91**	**-0.30**	**0.98**	**-0.22**

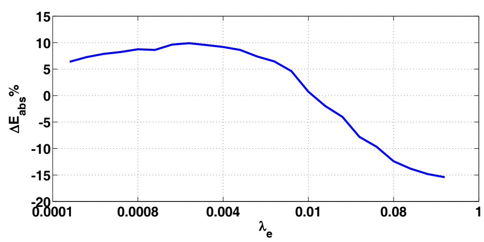
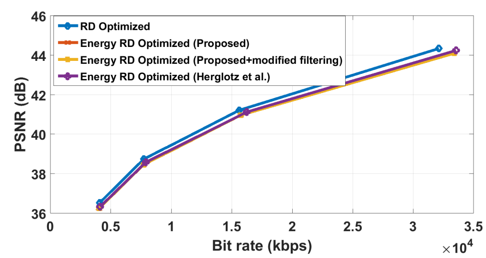


Fig. 4. The averaged variation of ΔE_{abs} with respect to λ_e for the set of training sequences in the Table 2.

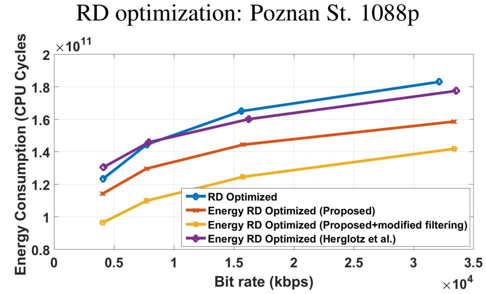
Intra Main configuration for 50 frames with QPs 22, 27, 32, and 37. Once more, the reference Intel core i7-2600 3.4 GHz CPU and 8GB RAM x86 system is used, together with the commonly used profiling tools [15], to evaluate the HM reference decoder's [17] energy consumption. The overall performance of the proposed energy-rate-distortion coding mode selection algorithm is presented in the Table 3. Moreover, Fig. 5 graphically illustrates the RD performance curves and the CPU cycle variations of the decoder for the different analysis scenarios discussed below.

Observing the RD performance in Fig. 5, it is evident that the coding efficiency reduction incurred due to non-RD optimized mode selection (Fig. 3), has been reduced significantly when using (10) as the optimization cost function, as opposed to using a purely energy based optimization as in (9). For example, the -1.618 dB BD-PSNR loss observed in Fig. 3 for the "Poznan Street" sequence, has been reduced to -0.25 dB when using the proposed energy-rate-distortion optimization. In addition, the generated bit streams have resulted in an average energy saving of 10.8% at the decoder with a minimal impact to the coding efficiency.

Furthermore, the ability to generate energy-efficient bit streams is tightly coupled with the accuracy and the level of detail of the underlying energy model utilized to predict the decoding energy of a CU. In this context, the effectiveness of the energy model proposed by Herglotz *et al.* [8] is also analyzed and compared with that of the proposed algorithm. Here, the mode selection in (10) is performed



RD optimization: Poznan St. 1088p



Energy saving: Poznan St. 1088p

Fig. 5. RD and energy saving performance of the proposed energy-rate-distortion optimized coding mode selection algorithm.

using [8] as the energy model that predicts the decoder energy consumption. It is observed that the average energy saving achieved is approximately -2 to 1% which is comparatively lower than the proposed method. The energy savings in Fig. 5 suggest that more energy is saved at lower QPs (higher bit rates), whereas it deteriorates as the QP increases (lower bit rates). Analysis reveals that this is due to the less accurate capturing of the variations in the decoding energy with respect to the TU hierarchy. Thus, the actual energy efficient modes within the optimization process are incorrectly identified, leading to less energy efficient bit streams being generated.

Finally, the performance of the proposed algorithm when combined with a dynamic manipulation of the in-loop filtering[3] process is evaluated. In this context, an algorithm based on [10] is implemented at the decoder to skip the in-loop filters,(i.e., de-blocking and SAO filters) based on a desired activation level [10]. The experimental results reveal that on average a 21.95% decoder energy saving can be achieved when the filters are skipped for all the frames, which is the highest activation level that can be achieved [10]. Furthermore, it is observed that the quality impact of the resultant operation is negligible compared to the scenario of having both in-loop filters enabled for all the encoded video frames. This suggests that the proposed decoder energy-aware intra-frame coding mechanism for HEVC can be extended to include in-loop filter manipulation to achieve an additional gain in the energy savings attained at the decoder.

In summary, the experimental results presented in this paper are based on the energy levels profiled for the HM reference decoder with respect to an Intel x86 reference system. Therefore, it is vital that the energy levels used within the model are adapted to cater for an arbitrary decoder which is not based on the HM decoder. However, the energy model and the optimization algorithm remains valid in general, hence, can be extended to an arbitrary system environment which runs a software decoder. Thus, the proposed algorithm has the potential to significantly improve the energy efficiency of the media streaming solutions.

6. CONCLUSION

A major conclusion to be drawn from this research is that the encoder could exploit the diversity of the decoder's energy requirements, in various coding modes, as an input parameter at the encoder, to generate decoder energy-aware video bit streams. In this context, the energy model proposed for HEVC intra-frame decoding, predicts the decoding energy of a CU with an average prediction error of less than 2%. Furthermore, the proposed energy-rate-distortion optimized coding mode and coding structure selection algorithm is capable of achieving an average decoding energy reduction of 10.81% with a BD-PSNR loss of -0.25 dB compared to the bit streams generated by the HM reference encoder. Moreover, the use of the proposed algorithm together with a method that skips the in-loop filtering, shows further reduction in the decoder's energy consumption. The future work will focus on extending the framework for HEVC inter-prediction and other system architectures.

7. REFERENCES

[1] Cisco, "Cisco visual networking index: global mobile data traffic forecast update 2014-2019," Cisco, White Paper.

[2] G.J.Sullivan, J. Ohm, W-J Han, and T. Wiegand, "Overview of the high efficiency video coding (HEVC) standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1649–1668, Dec. 2012.

[3] F. Bossen, B. Bross, S. Karsten, and D. Flynn, "HEVC complexity and implementation analysis," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1685–1696, Dec. 2012.

[4] Moving Picture Experts Group (MPEG), "Call for proposals on green MPEG," Apr 2013, Joint Collaborative Team on Video Coding (JCT-VC).

[3] In-loop filtering is a critical operation that is a significant energy consumer in the decoder. However, the filtering decisions are not typically included in RD optimization process when selecting the best coding modes, and are therefore considered to be outside the encoding loop.

[5] M. A. Hoque, M. Siekkinen, and J. K. Nurminen, "Energy efficient multimedia streaming to mobile devices a survey," *IEEE Commun. Surv. Tutorials*, vol. 16, no. 1, pp. 579–597, Feb 2014.

[6] F. Saab, I. H. Elhajj, A. Kayssi, and A. Chehab, "Profiling of HEVC encoder," *Electronics Letters.*, vol. 50, no. 15, pp. 1061–1063, July 2014.

[7] J. Henkel, M. U. K. Khan, and M. Shafique, "Energy-efficient multimedia systems for high efficiency video coding," in *Proc. IEEE International Symposium on Circuits and Systems (ISCAS)*, Lisbon, Portugal, May 2015, pp. 613–616.

[8] C. Herglotz, D. Springer, A. Eichenseer, and A. Kaup, "Modeling the energy consumption of HEVC intra decoding," in *Proc. IEEE International Conference on Systems, Signals, and Image Processing*, Bucharest, Romania, 2013, pp. 91–94.

[9] T. Mallikarachchi, H. Kodikara Arachchi, D. Talagala, and A. Fernando, "CTU level decoder energy consumption modelling for decoder energy-aware hevc encoding," in *Proc. IEEE International Conference on Consumer Electronics*, Las Vegas, USA, Jan 2016.

[10] E. Nogues, S. Holmbacka, M. Pelcat, D. Menard, and J. Lilius, "Power-aware HEVC decoding with tunable image quality," in *Proc. IEEE workshop on Signal Processing Systems.*, Belfast, Noth. Ireland, Oct 2014, pp. 1–6.

[11] H. Guo, K. Hu, and T. Xia, "Energy-efficient co-scheduling of receiving packets and decoding tasks on mobile video streaming terminals," in *Proc. IEEE International Conference on Multimedia and Expo Workshops.*, Chengdu, China, July 2014, pp. 1–6.

[12] W. Y. Liang, M. F. Chang, Y. L. Chen, , and C. F. Lai, "Energy efficient video decoding for the android operating system," in *Proc. IEEE International Conference on Consumer Electronics.*, Las Vegas, USA, Jan 2013, pp. 344–345.

[13] P. Shenoy and P. Radkov, "Proxy-assisted power-friendly streaming to mobile devices," in *Proc. Multimedia Computing and Networking Conference.*, 2003, pp. 177–191.

[14] Y. He, M. Kunstner, S. Gudumasu, E. S. Ryu, Y. Ye, and X. Xiu, "Power aware HEVC streaming for mobile," in *Proc. IEEE International Conference on Visual Communications and Image Processing.*, Kuching, Malaysia, Nov. 2013, pp. 2–6.

[15] Valgrind Developers, "The valgrind quick start guide," Valgrind Documentation.

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

[17] "HM 16.0," `https://hevc.hhi.fraunhofer.de/trac/hevc/browser/tags/HM-16.0`.

[18] J. Lainema, F. Bossen, W. J. Han, J. Min, and K. Ugur, "Intra coding of the HEVC standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1792–1801, Dec. 2012.

[19] G. Bjøntegaard, "Calculation of average PSNR differences between RD-curves.," *ITU - Telecommunications Standardization Sector STUDY GROUP 16 Video Coding Experts Group (VCEG)*, 2001.

[20] M. Lauridsen, "Power-consumption measurements for LTE user equipment application note," Jun. 2014, Agilent Technologies.