COMPLEXITY SCALABLE IDCT BASED APPROACHES FOR POWER-EFFICIENT VIDEO DECODING

Zahaib Akhtar, Mehar Ali, Nadeem A. Khan, Member IEEE, Jahangir Ikram, Member IEEE

Multimedia Research Lab
School of Science and Engineering, Lahore University of Management Sciences, Lahore, Pakistan

ABSTRACT

Running a multimedia application such as audio playback or video playback on a mobile handheld device is a power hungry task. Mobile devices in question here may include smart phones, palmtops or ipods. As the battery time is limited on these platforms, complexity cutting and power saving is a critical issue and optimization of such power intensive tasks is inevitable. This paper outlines two complexity scalable approaches for video decoding on embedded platforms.

***Index Terms*—**video decoding, low complexity, power efficient, mobile and embedded platforms

1. INTRODUCTION

With the rapid increase in semi-conductor technology during the last decade, the embedded processors industry has made great strides in developing high end processors for mobile and handheld devices. These gains have made the usage of processor intensive multimedia applications such as video decoding, recording and transmission on battery operated portable devices to be common tasks. In such devices power consumption is a critical issue. A battery-operated mobile device has only limited energy. Therefore, it is essential that power saving techniques must be implemented in video processing mobile applications. These techniques should be applied in algorithm, architecture and circuit levels. The main idea is to have a configurable design that has multiple power consumption modes.

Among the power consumption patterns of different applications on a typical mobile multimedia platform, video encoding and decoding are dominant in power consumption with video decoding being the more commonly performed task. Tang et al. and Wallach et al. have reported video decoding to occupy, on average, at least 30% of total power consumption [1][2]. The video decoding process of a macro-block is split up into various stages/modules, namely the Variable Length Decoding (VLD), Inverse Quantization (IQ), Inverse Discrete Cosine Transform (IDCT), Shape Decoding and Motion Compensation (MC). Of these tasks the most computationally intensive are IDCT and Shape Decoding, with IDCT taking up between 22-50% of the total power during the decoding process [3][4]. This paper proposes two complexity scalable IDCT based decoding approaches. The Complexity-Quality tradeoff is utilized such that significant decrease in computational complexity is achieved with a small decrease in the Peak Signal to Noise Ratio (PSNR). Low energy blocks are selected based on temporal and spatial redundant information in video sequences. The selected blocks then do not undergo IDCT thereby reducing the IDCT complexity by up-to 50-90% with a graceful decrease in the video PSNR.

2. RELATED WORK

As IDCT is the main power consuming module in the video decoding process, many techniques have been implemented to reduce IDCT computations and power consumption. Based on matrix decomposition, various fast IDCT algorithms have been proposed by researchers to reduce the computational complexity by decreasing number of multiplications and addition operations for an 8-point DCT/IDCT algorithm[5][6].

Some researchers have tried to optimize IDCT power consumption by classifying the blocks with respect to their energy level and pruning away those coefficients in the block which do not contribute much to the overall energy On the basis of this principle the blocks undergo IDCT with lesser number of coefficients resulting in complexity reduction [7][8]. Kim et al. [7] have made this classification in three categories depending on the energy levels of macroblock coefficients. IDCT block sizes of 1x1, 3x3 or 4x4 are then selected from among the 8x8 coefficient blocks. All the remaining coefficients are taken as zero valued coefficients.

Similarly, Peng [8] has devised an algorithm which assigns a higher priority to the lower frequency components in the upper left corner of the DCT coefficient matrix. As the coefficients in the upper left corner contain higher energy the data is pruned from the lower right corner which gives minimal quality degradation.

Choi et al. have proposed a zero-coefficient aware algorithm which selects between conventional butterfly IDCT algorithm and zero skipping IDCT algorithm [9]. The decision is made on the basis of non-zero coefficients present in a particular Macroblock, resulting in reduction of the IQ and IDCT computation.

3. APPRAOCHES EXPLORED FOR REDUCTION IN IDCT COMPUTATIONS

A frame consisting of Macroblocks (MB) offers various avenues for the reduction of complexity at the cost of inducing distortion in the video sequence. The inter-coded mechanism of video sequences incorporates MBs in a frame which may entirely consist of zero valued coefficients, whereas there may be some MBs which have a few non-zero coefficients and overall do not contribute much to the perceptual quality of the frame image. These MBs may be characterized as low energy blocks. Identifying them in a video frame and treating them as an all-zero coefficient block can potentially result in lesser complexity as the number of computationally complex IDCT operations will decrease. However, as stated above, this optimization will raise the distortion level in the video. Our task then is to design the selection algorithm by keeping the Complexity-Quality tradeoff in consideration, such that only those blocks which do not significantly contribute to the perceptual quality may be targeted, thereby decreasing the complexity without significantly affecting the PSNR.

3.1. Re-quantization Based Approach

One of the approaches explored was to detect the low energy blocks by performing Coarse Re-Quantization (CRQ) on the coefficients of the Macroblocks. The steps were performed in the Inverse Quantization (IQ) stage of the decoding process. Just after the coefficients have gone through IQ, they are re-quantized by a CRQ value greater than the default Quantization Parameter (QP) used at video encoding.

$$CRQ = QP^* \text{ where } QP^* > QP \text{ and } 16 < QP^* \leq 72$$

The default value used for QP is 16 and QP* denotes the new quantization parameter which will be used as the CRQ value. As the CRQ value is greater than the default quantization parameter a lot of low value coefficients may become zero after re-quantization. This leads some blocks (which previously had a few low valued non-zero coefficients) to now only possess zero-value coefficients. Such blocks are then flagged as zero coefficient blocks and do not undergo IDCT operation during the decoding process resulting in complexity reduction due to a decrease in the number of computations.

3.2. Coefficient Absolute Sum Based Approach

In this approach, the detection of low energy blocks was done by computing the Coefficient Absolute Sum (CAS) of non -zero coefficients from a Macroblock. The principle idea behind this approach was that low energy blocks

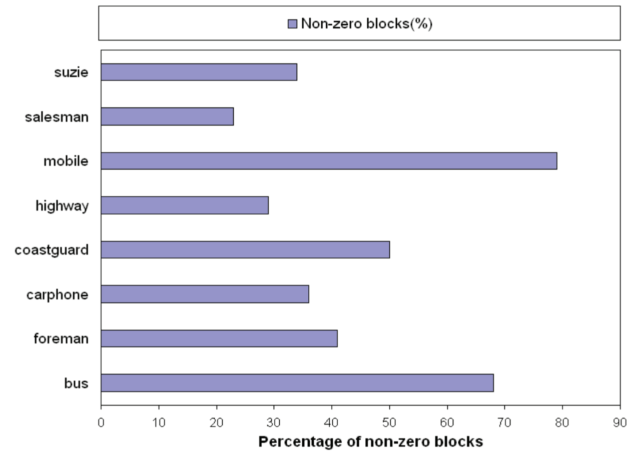

Fig 4.1 Percentage of blocks with atleast one non-zero coefficient for different QCIF (176×144) video sequences

would comprise of a majority of zero coefficients and a few non zero coefficients. Therefore, the absolute sum of coefficient values for such blocks will also be a small value. This sum was calculated and then tested against a threshold. The range of threshold varied from 0 to 30, with 0 being the base threshold value.

$$CAS = \sum_{i=0}^{63} |C_i|$$

If the absolute sum was greater than the threshold, the block was considered to be a high energy block and decoded using the normal video decoding pipeline. On the other hand, if the absolute sum of the block coefficients is found to be less than the desired threshold, a flag signifying low energy is placed and all non-zero coefficients of this block are then replaced by zeros. These blocks do not undergo IDCT as performing IDCT on all-zero valued coefficients will be of no significance.

4. RESULTS AND DISCUSSION

Experiments to find out the percentage of all-zero coefficient blocks were conducted on a number of video sequences. As expected the percentage of all-zero coefficient blocks was higher in low motion video sequences, whereas for high motion videos the percentage was quite low as demonstrated by "bus", "mobile" and "coastguard" sequences in Fig. 4.1. From this result it can be concluded that the scheme proposed by Choi et al. would not be a very good optimization scheme for high motion videos as the optimization achieved will not be very significant due to low percentage of all-zero coefficient blocks. The scheme also does not exploit the spatial and temporal redundancy inherent in video frames by targeting

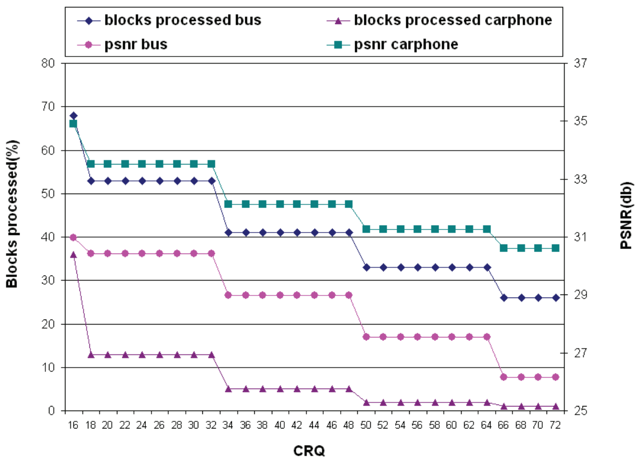

Fig 4.2 **Blocks processed and corresponding PSNR against CRQ parameter using "bus" and "carphone" sequence.**

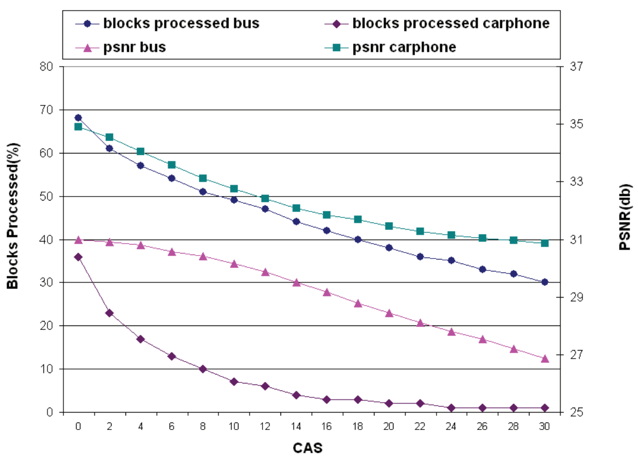

Fig 4.3 **Blocks processed and corresponding PSNR against CAS parameter using "bus" and "carphone" sequence**

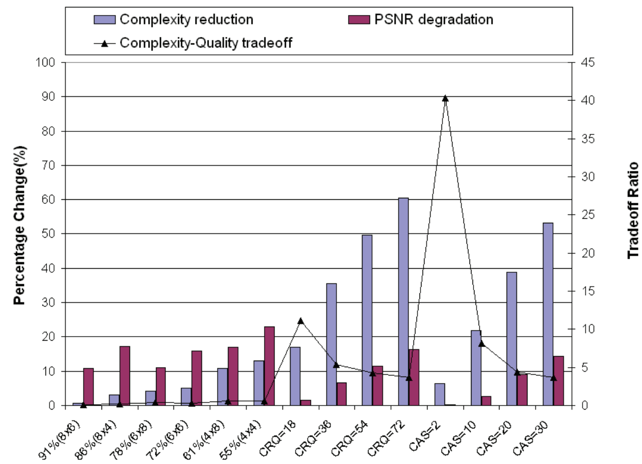

Fig 4.4 **Percentage drop in complexity and Complexity-Quality tradeoff for the proposed approaches using "bus" QCIF sequence.**

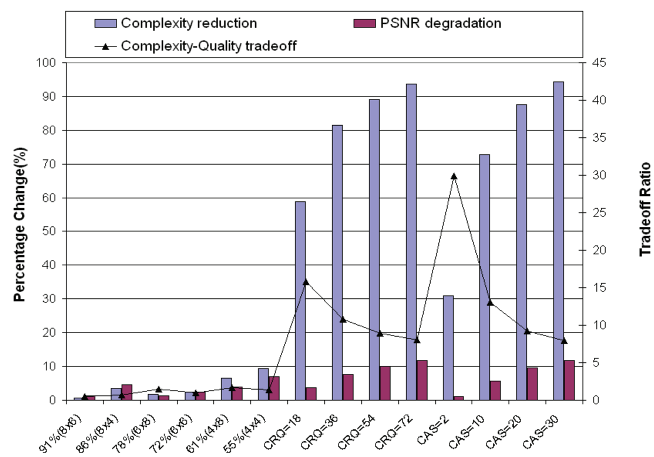

Fig 4.5 **Percentage drop in complexity and Complexity-Quality tradeoff for the proposed schemes using "carphone" QCIF sequence.**

those blocks which only have a very few non-zero coefficients.

The proposed approaches were applied to (176×144) sized QCIF video sequences which consisted of I- and P-frames. The above discussed approaches were only applied on P-frames. Simulations were conducted on "bus" and "carphone" video sequences. The complexity was analyzed by measuring IDCT clock cycles using Sim- Wattch[10]. As the IDCT module is the most complex and power intensive process during video decoding, optimization in it results in significant complexity reduction of the decoder, thereby reducing the power consumption as well. The percentage of blocks processed at different CRQ and CAS levels along with the corresponding PSNR were recorded. Figure 4.2 shows the decrease in the percentage of blocks processed and the drop in the PSNR at different CRQ parameter levels for the "bus" and "carphone" sequences. It can be seen from the results that the decrease in the percentage points of the processed blocks for "carphone" sequence is approximately 34 and approximately 90% of the non-zero blocks which were processed at the base value, are not processed at the highest CRQ value. Similarly, the decrease in percentage points of the processed blocks for "bus" is about 42 and about 62% of the non-zero blocks are not processed at the highest CRQ value. On the other hand, it can also be seen that the drop in PSNR is much more graceful. The drop in PSNR (in dB) for "bus" QCIF sequence is 4.83dB (15.55% drop) whereas, the drop in PSNR (in dB) for "carphone" is about 4.82dB (13.9% drop). Almost similar trend in the percentage of blocks and corresponding PSNR can be observed from Fig 4.3.

To do a comparative analysis, the scheme proposed by Peng et al. was implemented in the decoder and its computational complexity was calculated by measuring IDCT clock cycles. Fig 4.4 and 4.5 present the results of the proposed approaches in comparison to the scheme proposed by Peng et al. It can be observed that the optimizations

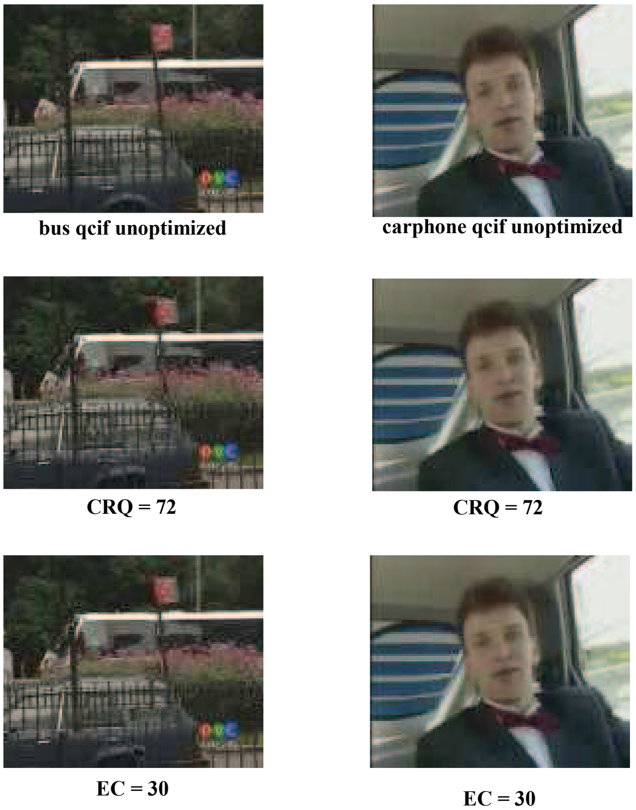

Fig 4.6 Decoded video frames at various levels of optimizations using bus and carphone QCIF sequences

achieved with the proposed approaches present higher complexity reduction with smaller decrease in the PSNR. The Quality-Complexity tradeoff is on the average 4 times higher than that achieved by Peng et al. A few screen shots from the videos at extreme CRQ and CAS parameter show that the PSNR degraded gracefully even at the highest levels of optimizations.

Comparison among the proposed approaches shows that the CAS approach presents much finer control over the drop in PSNR and percentage of blocks process. As can be seen from Fig. 4.3 that there is a much more smooth and uniform drop than the step like drop shown in Fig 4.2. This variation in behavior is due to the difference inherent in both approaches. The Re-quantization based approach treats coefficient divison and multiplication as integer operations, therefore step like graphs are seen instead of smooth curves which are obtained from the CAS based approach. Generally both the approaches bring about similar decrease in the percentage of blocks processed with approximately equal drop in PSNR. However, as can be seen from Fig 4.4 and 4.5, the CAS approach performs better than the CRQ approach on the Quality-Complexity tradeoff at initial levels of optimization. At higher optimization levels both the approaches exhibit approximately similar tradeoff.

5. CONCLUSION

In this paper we have studied and proposed two approaches to cut IDCT complexity and therefore reduce power consumption in a video decoder. The influence of both the approaches on Quality-Complexity tradeoff has been analyzed. Both have shown efficient results, cutting down IDCT complexity by 50-90% with a graceful 10-15% drop in PSNR.

6. REFERENCES

[1] Wen Ji, Yiqiang Chen, Chen Lei, Huigo Zhang, Hui Wang, Yunbing Xing, Bo Tang, *"Module and Distortion Analysis for Video Decoding on Mobile Devices"*, Seventh International Conference on Networking

[2] Viredaz, M. A. and Wallach, D. A. Power evaluation of a handheld computer: a case study. *Tech. Rep, Compaq Western Research Laboratory, May 2001*

[3] Fabio Cavalli, Rita Cucchiara, Massimo Piccardi, Andrea Prati, Performance Analysis of MPEG-4 Decoder and Encoder, VIPromCom-2002, 4th EURASIP

[4] H.C. Chang, L.G. Chen, M.Y. Hsu, Y.C. Chang, Performance analysis and architecture evaluation of MPEG-4 video codec system, IEEE Int. Symp. on Circuits and Systems,2000, vol. II, 449-452

[5]B. G. Lee, "A new algorithm to compute the discrete cosine transform," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-32, pp.1243–1245, Dec. 1984.

[6]H. S. Hou, "A fast recursive algorithm for computing the discrete cosine transform," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-35, pp. 1455–1461, Oct. 1987.

[7]D. Kim and D. Shin "Energy-based adaptive DCT/IDCT for video coding" Multimedia and Expo, 2003. ICME '03. Proceedings July 2003.

[8]S.Peng "Complexity scalable video decoding via IDCT data pruning". Consumer Electronics, 2001. ICCE. International Conference on June 2001.

[9]K.Choi, K. Lee, E. Kim and E. S. Jang, "Zero Coefficient-aware fast IQ-IDCT Algorithm " Network Infrastructure and Digital Content,IEEE International Conference Vol. 56, No. 3, August 2010

[10]D. Brooks, V. Tiwari, M. Martnosi "Wattch: A Framework for Architecture-Level Power Analysis and Optimizations" Proceedings of the 27th International Symposium on Computer Architecture, June 2000