

Complexity scalable motion estimation control for H.264/AVC

Citation for published version (APA):

Huijbers, E. A. M., Ozcelebi, T., & Bril, R. J. (2011). Complexity scalable motion estimation control for H.264/AVC. In *29th International Conference on Consumer Electronics (ICCE 2011, Las Vegas NV, USA, January 9-12, 2011)* (pp. 49-50). Institute of Electrical and Electronics Engineers. <https://doi.org/10.1109/ICCE.2011.5722705>

DOI:

[10.1109/ICCE.2011.5722705](https://doi.org/10.1109/ICCE.2011.5722705)

Document status and date:

Published: 01/01/2011

Document Version:

Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy

If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

Complexity Scalable Motion Estimation Control for H.264/AVC

E. (Rico) A. M. Huijbers, Tanır Özçelebi, Reinder J. Bril

Department of Mathematics and Computer Science, Eindhoven University of Technology
P.O. Box 513, 5600 MB Eindhoven, The Netherlands

Abstract— **Guaranteeing real-time performance for video encoding on platforms with limited resources is becoming increasingly important for consumer electronics applications. In this paper, an extension of an H.264/AVC encoder with complexity scalable motion estimation (ME) control is presented. An upper bound on the complexity of encoding a single frame is achieved by restricting Sum of Absolute Differences (SAD) computations performed during ME and trading complexity allocation per frame for output quality. Allocation based on residual, i.e. SAD distortion of the final ME match before quantization, of the co-located macroblock in the previous frame outperforms other approaches in the literature in video quality.**

I. INTRODUCTION

Video encoding has computational requirements that vary over time based on the video content. It is difficult to guarantee real-time performance on a platform where computational resources are limited or fluctuating. A solution is to use a complexity scalable video encoder that can adapt its workload and stay within a computational budget. Apart from allowing real-time operation with reduced energy consumption, such an encoder is reusable in various systems, saving consumer electronics manufacturers time and money.

We study an H.264/AVC [1] encoder running at a fixed bitrate on a resource constrained platform and in real-time, i.e. the video frames must be produced strictly periodically. The encoding complexity, measured in clock cycles, can be arbitrarily restricted at the frame level via scaling motion estimation (ME); i.e. restricting the number of Sum of Absolute Differences (SAD) computations per frame. Finally, *residual-based* resource allocation is performed on the macroblock (MB) level as explained in Section 4.

II. RELATED WORK

There are related works in the literature that base their MB coding complexity allocation decisions solely on the statistics of collocated MBs in the previous frames, e.g. [2], [3]. In [2], rate-distortion (RD) performance of encoder is modeled based on the scalability of frame rate, and number of SAD operations and DCT computations. The approach of [3] is based on the history of RD gain of the collocated MB in the previous frame, while we propose decisions based on difficulty of finding a ME match for this MB. In [4], the encoder is restricted to a set of ME operating modes, as opposed to the encoder presented here, which operates across the entire range from min to max complexity. In [5], complexity scalability is achieved by varying the number of ME operations at the frame level and DCT computations at the MB level. The drawback is that access to all video frames is required before encoding begins.

III. COMPLEXITY SCALABLE ME

ME is computationally very intensive, taking up to 60-80% of total encoding time [6]. Therefore, our focus is on making complexity scalable ME. The SAD calculation dominates the computational cost of ME. By restricting the number of motion vectors (MV) examined, we introduce a *complexity budget* for every MB that indicates the number of SAD computations allowed and we modify the search algorithm to obey this budget. The fixed complexity cost of a single SAD computation is used to convert the unit of complexity budget into number of clock cycles. The cost of an SAD computation is a linear function of the number of pixels examined, i.e., the area of the block size that is checked. We select the examination of a single 16x16 MB as the unit of our complexity budget and every SAD calculation of mode μ (e.g. $\mu:8x4$) consumes an MB complexity budget of $(w_\mu \cdot h_\mu)/(16 \cdot 16)$, where w_μ and h_μ are the width and height of a block in mode μ . The complexity cost of a frame of M MBs, every MB having a budget of B_m ($0 \leq m < M$) can now be expressed as:

$$C_{frame} = \alpha \cdot B_{frame} + \beta \cdot M + \gamma \quad (1)$$

$$B_{frame} = \sum_{0 \leq m < M} B_m \quad (2)$$

where B_{frame} is the complexity in budget units and C_{frame} is the complexity in clock cycles. α is the complexity cost of one SAD calculation, β is the cost of non-scalable operations associated with every MB and γ is the overhead cost of processing a frame (all in clock cycles).

IV. FRAME BUDGET ALLOCATION

We can use the model presented in the previous section to find a frame budget B_{frame} given a desired complexity bound in clock cycles C_{frame} . Then, the frame budget is partitioned into MB budgets B_m , in such a way that the quality loss incurred from doing reduced computation is minimized. One allocation strategy is to divide the frame budget equally among all MBs, i.e. *uniform allocation*. It is shown in [2] that the *Motion History Matrix (MHM)* allocation, which uses the MVs found during ME to estimate the complexity needs, outperforms uniform allocation. Our goal is to identify the MBs that will benefit the most from extra steps of ME and allocate the complexity budget to them. We use information gathered during the encoding of previous frames to estimate the complexity needs of MBs in the current frame. For this, we use the distortion of the co-located MB encoded in the previous frame as the indicator. A high distortion indicates a need and a possibility for improvement: if an MB had a high distortion in the previous frame, this is an opportunity to

invest more computational power into it, with the goal to find a better ME match and to decrease the achieved distortion. We identify 2 strategies, i.e. *Distortion History (DH)* and *residual* allocation.

DH for an MB is defined as the SAD achieved while encoding the co-located MB in the previous frame. However, this metric depends on ME as well as quantization. We propose *residual*-based allocation, which is the distortion of the prediction error without quantization, i.e. it is calculated *before the quantization stage of encoding*. The residual is the distortion between the macroblock in the previous frame, and the best motion vector candidate in the reference frame before that. It is a measure of how hard it is to encode an MB, and specifically, to find a good ME match for an MB.

To prevent wastage of leftover budget from previous MBs (e.g. when a good enough MV is found immediately), we re-evaluate the allocation at the start of every MB using the remaining frame budget, B'_m . B_m is a fraction of B'_m as given by (3) where x_i is either the DH or the residual of MB i .

$$B_m = \frac{x_m}{\sum_{m \leq i < M} x_i} \cdot B'_{frame} \quad (3)$$

V. RESULTS AND CONCLUSION

We have implemented the complexity scalable ME in the reference encoder JM 14.2. We encoded CIF test video sequences at various complexity bounds and measured the complexity in clock cycles per frame. We established a linear relationship between complexity budget in number of SAD operations and actual complexity in clock cycles, as shown in Fig. 1 for the “foreman” sequence. Since the experiments were performed on a time-sharing operating system, there is noise in the graph where the encoder was pre-empted.

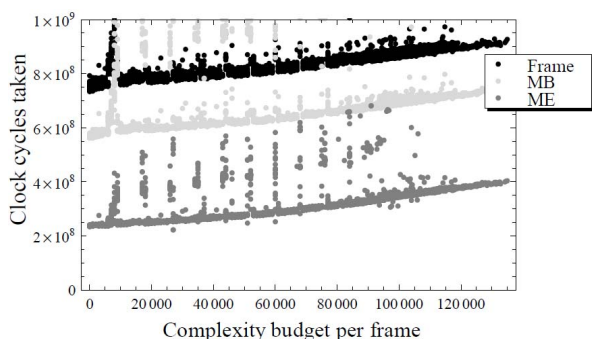


Figure 1: Actual complexity of encoding, measured in clock cycles, as a function of the complexity budget B_{frame} . The three point clouds represent total frame, total MB and ME complexity.

To evaluate the complexity allocation algorithms, a variety of CIF videos were encoded at a fixed rate using one reference frame, full-pixel ME, and low complexity mode decision algorithms from JM 14.2. A range of complexity budgets (per frame) were employed; using complexity allocation algorithms based on each of the two metrics (DH and residual). In every encoding, all frames have the same complexity bound and leftover frame budget is not reused. In all cases the residual allocation yielded higher quality video.

Fig. 2 shows a plot of the complexity budget against the achieved quality for the CIF sequence “bus” encoded at 1500 kbps.

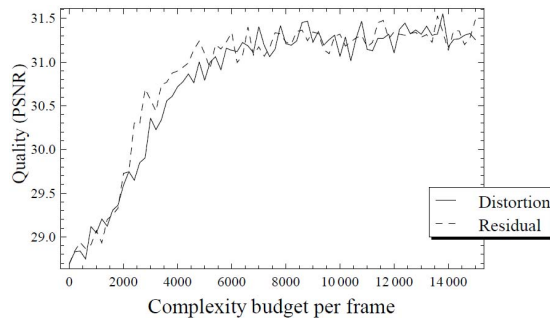


Figure 2: PSNR for DH versus residual based allocation.

In Fig. 3, we compare residual based allocation with *uniform* and *MHM* allocations of [2]. Note that the proposed residual-based scheme results in a PSNR gain of up to 1 dB below a complexity budget of 6000 SAD computations per frame, and yields comparable PSNR with the MHM allocation above this complexity budget.

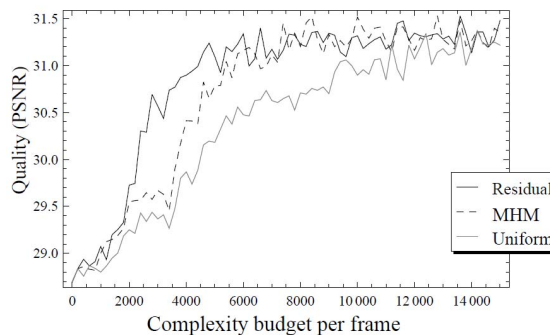


Figure 3: Uniform, MHM and residual-based allocation.

To summarize, we extended the work done in [2] to the domain of H.264/AVC encoding. We showed that using the number of SAD computations as a budget allows us to bound the complexity of encoding to an arbitrary amount of clock cycles. We investigated a number of budget allocation strategies. The residual based allocation predicts the complexity needs for MBs with better accuracy in all cases.

REFERENCES

- [1] T. Wiegand, G.J. Sullivan, G. Bjontegaard, and A. Luthra, “Overview of the H. 264/AVC video coding standard,” *IEEE Trans. on Circ. and Sys. for Vid. Tech.*, vol. 13, no. 7, pp. 560–576, 2003.
- [2] 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 Circ. and Sys. for Vid. Tech.*, vol. 15, no. 5, pp. 645–658, 2005.
- [3] C. Kim, J. Xin, A. Vetro, and C.C. Jay Kuo, “Complexity Scalable Motion Estimation for H.264/AVC”, *SPIE Conference Visual Communications and Image Processing*, Jan 2006.
- [4] L. Su, Y. Lu, F. Wu, S. Li, and W. Gao, “Real-time video coding under power constraint based on H.264 codec,” *Proc. of SPIE, Visual Comm. and Image Proc.*, vol. 6508, pp. 650802, 2007.
- [5] S. Mientens, P. H. N. de With, and C. Hentschel, “New complexity scalable MPEG encoding techniques for mobile applications,” *EURASIP J. Appl. Signal Process.*, vol. 2004, pp. 236–252, 2004.
- [6] H. F. Ates and Y. Altunbasak, “Rate-Distortion and complexity optimized motion estimation for H.264 video coding,” *IEEE Trans. Circ. Syst. Video Technology*, 18, 2, pp. 159–171, 2008.