



3D high definition video coding on a GPU-based heterogeneous system [☆]



Rafael Rodríguez-Sánchez ^{a,*}, José Luis Martínez ^a, Jan De Cock ^c, Gerardo Fernández-Escribano ^a, Bart Pieters ^c, José L. Sánchez ^a, José M. Claver ^b, Rik Van de Walle ^c

^a Instituto de Investigación en Informática de Albacete, Universidad de Castilla-La Mancha, Avenida de España s/n, 02071 Albacete, Spain

^b Departamento de Informática, Universidad de Valencia, Avenida de la Universitat s/n, 46100 Burjassot, Valencia, Spain

^c Multimedia Lab, Ghent University-IBBT, Gaston Crommenlaan 8 bus 201, B-9050 Ledeberg-Ghent, Belgium

ARTICLE INFO

Article history:

Available online 20 June 2013

ABSTRACT

H.264/MVC is a standard for supporting the sensation of 3D, based on coding from 2 (stereo) to N views. H.264/MVC adopts many coding options inherited from single view H.264/AVC, and thus its complexity is even higher, mainly because the number of processing views is higher. In this manuscript, we aim at an efficient parallelization of the most computationally intensive video encoding module for stereo sequences. In particular, inter prediction and its collaborative execution on a heterogeneous platform. The proposal is based on an efficient dynamic load balancing algorithm and on breaking encoding dependencies. Experimental results demonstrate the proposed algorithm's ability to reduce the encoding time for different stereo high definition sequences. Speed-up values of up to 90× were obtained when compared with the reference encoder on the same platform. Moreover, the proposed algorithm also provides a more energy-efficient approach and hence requires less energy than the sequential reference algorithm.

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

The H.264/MPEG-4 Advanced Video Coding (AVC) standard [1] and its successor, High Efficiency Video Coding (HEVC) [2], have demonstrated significant improvements in video compression capabilities in the past few years [1], and currently [2], for a wide range of bit-rates and resolutions [3]. Moreover, since developing the H.264/AVC standard, both the Joint Video Team (JVT) of the ITU-T Video Coding Experts Group (VCEG) and the ISO/IEC Moving Picture Experts Group (MPEG) have also standardized an extension of it, which is referred to as Multi-view Video Coding (MVC) [4]. The Joint Collaborative Team on 3D Video (JCT-3V) is currently working on the 3D extension of HEVC, and on the corresponding 3DV-HTM software. Therefore, the one standard solution for multi-view/stereo video coding at the moment is the H.264/MVC implementation.

H.264/MVC can provide users with the impression of complete scene perception by simultaneously transmitting several views to the receivers. It can also give users vivid information about the scene structure. H.264/MVC considers not only the correlation in the temporal direction for motion-compensated prediction (MCP), but also in the view direction for disparity compensated prediction (DCP). MVC adopts many coding tools used for single view H.264/AVC, such as the variable block-size Motion Estimation (ME) algorithm (which will be described in Section 2), among others. This is a very complex algorithm and, as part of the whole inter prediction algorithm, it is applied many times in terms of different MacroBlock (MB)

[☆] Reviews processed and recommended for publication to Editor-in-Chief by Guest Editor Prof. Jesus Carretero.

* Corresponding author. Tel.: +34 967599200x2663.

E-mail addresses: rrsanchez@dsi.uclm.es (R. Rodríguez-Sánchez), joseluismm@dsi.uclm.es (J.L. Martínez), jan.decock@ugent.be (J. De Cock), gerardo@dsi.uclm.es (G. Fernández-Escribano), bart.pieters@ugent.be (B. Pieters), jsanchez@dsi.uclm.es (J.L. Sánchez), jose.claver@uv.es (J.M. Claver), rik.vandewalle@ugent.be (R. Van de Walle).

sizes and, in the particular case of multi-view, between different views. All these computations involve even higher encoder complexity than single view H.264/AVC, which makes it necessary to develop a method that can reduce the complexity of MVC with minimal loss in coding efficiency (bit-rate and quality). In this area, several approaches are available in the literature that aim to accelerate this process by using algorithmic and code optimizations. However, as far as the authors of this paper know, there are no approaches that make use of hardware platforms to achieve this goal.

On the hardware side, in the past few years new architectures have been introduced in high-performance computing. Examples of such architectures include systems composed of multi-core CPUs and Graphics Processing Units (GPUs). The heterogeneous nature of modern desktop systems imposes additional challenges for module parallelization. Several programming frameworks, such as the CUDA (Compute Unified Device Architecture) [5] programming model, already address the programmability issues in CPU + GPU environments by offering the unification of the execution space for different heterogeneous devices. CUDA abstracts both Simple Instruction Multiple Data (SIMD) and task parallelism into thousands of simultaneous threads.

At this point, this paper proposes an algorithm that efficiently distributes the burden of complexity of the inter prediction algorithm executed in H.264/MVC Stereo High Profile on a GPU-based heterogeneous platform. This paper is an extension of the one previously presented in [6]. This extension includes a more accurate and complex way of obtaining the motion information, surpassing the Rate Distortion (RD) performance obtained in [6]. Moreover, it includes a more complete evaluation section, as the proposals are tested against two of the most well-known inter prediction algorithms implemented in the reference software: the full search and UMHexagonS algorithms [7].

The algorithm includes both pixel and sub-pixel accuracy, as well as temporal and inter-view predictions and different transform sizes. Although the algorithm described in this paper is evaluated for stereo (2 views), it could easily be extended to n -views, and it can also provide the capability of 3D perception. The idea of the proposed algorithm consists in starting with the smallest MB sub-partitions, and is able to build the entire tree-structured ME procedure from bottom to top. In addition, some structural references have to be broken or adapted when they are calculated in parallel. Performance evaluation is carried out for full High Definition (HD) stereo video sequences and adapted for a system composed of an NVIDIA GTX480 and an Intel Core i7 processor. The results show a noteworthy time reduction of up to 98% with only a negligible RD penalty. Moreover, this work also shows that the GPU-based H.264/MVC encoding algorithm consumes less energy than the baseline algorithms running on a CPU core. Additionally, the proposed algorithm is tested against full and UMHexagonS search algorithms. GPU-based platforms raise the power consumption but the execution time is much shorter, which leads to a more efficient use of energy.

There are many potential applications for the algorithms presented in this paper. Nowadays, GPUs are available in a wide variety of environments, ranging from small and cheap personal computers, to large and expensive supercomputers. GPU manufacturers provide different GPU solutions to satisfy the requirements of all these different environments. As a consequence, the solutions proposed in this paper can be used by personal computers when performing a video conference, by more powerful servers dedicated to video streaming or by large high performance computers in the video storage industry.

The rest of the paper is organized as follows. Section 2 outlines the technical background. In Section 3 some related work is presented. Section 4 describes the approach presented in this paper. In Section 5 the proposal presented is evaluated. Finally, conclusions are given in Section 6.

2. Background

2.1. H.264/MVC

The H.264/MVC coding standard was recently finalized by JVT, and was developed as an extension of H.264/AVC. MVC mainly uses H.264/AVC while taking advantage of temporal and inter-view dependencies [8] which are based on Hierarchical Bidirectional Pictures prediction to exploit both temporal and inter-view correlations [9]. Moreover, MVC provides other techniques such as Disparity Estimation (DE), which is used in the process of inter-view coding.

H.264/MVC inherits many video coding techniques from H.264/AVC, such as multiple reference frames, weighted prediction, a de-blocking filter, variable block size and quarter-pixel precision for motion compensation, which allow this optimum performance to be achieved at the expense of an increase in the computational complexity of the encoder.

Basically, variable block-size matching ME [3] and DE [8] are used to reduce the temporal and inter-view redundancies between frames. In this coding system, variable block-size ME is carried out using eight inter prediction modes (SKIP, Inter 16×16 , Inter 16×8 , Inter 8×16 , Inter 8×8 , Inter 8×4 , Inter 4×8 , and Inter 4×4), which are depicted in Fig. 1. MVC determines which partitions deal with cost as the best MB partition. This results in a high complexity encoder. Therefore, it is necessary to develop a method that can reduce the execution time of MVC with minimal loss of image quality.

2.2. Graphics Processing Units (GPUs)

Recent heterogeneous platforms include GPUs in order to achieve high-performance computing. Although GPUs come primarily from interactive applications such as multimedia and computer or console gaming, they can be used for running

Download English Version:

<https://daneshyari.com/en/article/455657>

Download Persian Version:

<https://daneshyari.com/article/455657>

[Daneshyari.com](https://daneshyari.com)