早稲田大学大学院情報生産システム研究科

博 士 論 文 概 要

論 文 題 目

Research on Architecture Design for Ultra High Definition Video Encoding

申 請 者

Gang HE

情報生産システム工学専攻
情報生産システム研究

2013 年 11 月

Ultra high definition (UHD) video includes 4K (3840×2160) and 8K (7680×4320, also known as super Hi-vision) formats. By delivering 4 and 16 times of pixels per frame compared to today's high definition (1920x1080), UHD videos offers remarkably enhanced visual experience and provides rich cues for stereoscopic feeling such as visual field angle, linear perspective, and texture gradient. The UHD is currently being promoted in the next-generation standard of digital television.

To store and transmit the huge volume of UHD video data, efficient and real-time compression is essential. The latest video coding standards such as H.264/AVC and H.265/HEVC provide excellent compression ratio. However, the key compression algorithms, such as intra prediction and fractional motion estimation (FME), involve high computational complexity. Moreover, due to data dependency and the algorithm process, it is limited to apply the pipelining and parallel processing techniques in hardware design. The existing works proposed many architecture designs of intra prediction and FME, mainly targeting for 1080p HD or lower resolution videos. We cannot realize the UHD-throughput design by simply increasing times of previous architectures, since it leads to incredibly large hardware cost and high power for a chip.

Therefore, the target of this dissertation is to develop the efficient architectures of intra prediction and FME in H.264 or HEVC, targeting for UHD videos. Firstly, for intra prediction, an MB/Block co-reordering method is proposed. By doing so, we can parallelize the mode decision and reconstruction processes, and hardware utilization is improved by almost 50%. An H.264 VLSI architecture is designed for 4kx2k UHD. Secondly, to achieve higher throughput and reduce more complexity, an interlaced block reordering strategy, together with a preliminary mode decision are proposed. The hardware reusing methods are also used to optimize the architecture. As a result, the hardware complexity (the product of hardware area and required operating frequency) is reduced by 77%. The design achieves not only the throughput of 8kx4k UHD, but also high hardware utilization. Thirdly, a high-performance and low-power HEVC FME architecture is designed, with the co-optimization in algorithm and hardware architecture. An approximation interpolation, together with a directional search pattern reduces the complexity. The exhaustive-size hadamard transform is adopted to improve coding quality. The design realizes the real-time encoding for 8kx4k UHD and achieves much improvement on power efficiency.

The dissertation consists of 5 chapters as follows:

Chapter 1 [Introduction] gives a brief introduction to ultra-high definition video and video compression system. The basic knowledge of intra prediction and

fractional motion estimation are introduced, followed by the contributions and an overview of this dissertation.

Chapter 2 [An H.264 Intra Prediction Architecture for 4kx2k UHD] first introduces the design challenges of intra prediction, and then presents the proposed architecture based on H.264 for 4kx2k UHD.

Due to the high data dependency of intra prediction in H.264, both pipelining and parallel processing techniques are limited to be applied. Moreover, it is difficult to get high hardware utilization and throughput because of the long MB-level and block-level reconstruction loops.

Many researchers has proposed the hardware architectures of intra prediction. Lin (TVLSI'2009) proposed a interlaced schedule for prediction modes of different sizes to improve the pipeline efficiency, but it is only limited to the mode decision component. In order to eliminate the data dependency, Mochizuki (JSSC'2008) and Chuang (PCS'2007) used original pixels instead of reconstructed one to do the prediction, but brings large quality loss. Moreover, the above designs targets at HD or lower resolutions.

In this chapter, an H.264/AVC intra prediction architecture for 4kx2k UHD is presented. The proposed macroblock (MB) and block co-reordering can avoid data dependency and improve pipeline utilization by almost 50%. The timing constraint of real-time 4096x2160 encoding can be achieved with negligible quality loss. 16x16 and 8x8 prediction engines work parallel for generating coefficients. A reordering interlaced reconstruction is also designed for fully pipelined architecture. It takes only 160 cycles to process one MB. Hardware utilization of the prediction and reconstruction modules achieve 90-100%. Furthermore, PE-reusable 8x8 intra predictor and hybrid SAD & SATD mode decision are proposed to save hardware cost. The design is implemented by 90nm CMOS technology with 113.2k gates and can encode 4096x2160 video sequences at 60 fps with operation frequency of 332MHz. Compared with previous designs of 1080p@30fps (Lin,TCVST'2009, Chuang,PCS'2007), throughput of the design increases for 8 times with the overhead of less than 30% hardware cost.

Chapter 3 [An H.264 Intra Prediction Architecture for 8kx4k UHD] discusses the design problems for very high throughput and presents an H.264 8kx4k UHD intra-prediction architecture.

Firstly, due to this huge throughput requirement, design challenges such as complexity and data dependency, which currently exist for lower resolutions (e.g. 2160p and 1080p), become even more critical. Moreover, Pipeline latency influences the efficiency of pipelines with serious data dependency.

In this work, we first propose an interlaced block reordering scheme together

with a preliminary mode decision (PMD) strategy to resolve the data dependency between intra mode decision and reconstruction. In the meantime, hardware cost is reduced by PMD. We also propose a probability-based reconstruction scheme to solve the problem of long pipeline latency. In addition, hardware reuse strategies including a shared fine decision module and processing element-reusable prediction generator, are applied to further optimize the design. As a result, the hardware complexity (the product of hardware area and required operating frequency) is reduced by 77%, and it takes an average of 33 cycles to process an MB. The implementation result demonstrates that our design can support up to the specification of 7680 × 4320p 60 f/s when running at 273 MHz. The 1080p 30 f/s encoding requires less than 9 MHz operating frequency, which is much lower than those (114MHz, 140MHz) used in previous works (Kuo:TVLSI'2011, Lin:TCSVT'2009).

Chapter 4 [An HEVC Fractional Motion Estimation Architecture for 8kx4k UHD] presents a high-performance low-power HEVC FME architecture design.

FME improves the rate-distortion performance significantly about 3-6dB, but results in high computation complexity (40% encoding time) due to the complex interpolation and fractional search process. Previous works have proposed efficient designs in H.264 (Tsung,ICME'2009, Kao,TVSLI'2010). However, since HEVC involves many FME-related new features, such as 8-tap interpolation and quad-tree coding structure, they cannot be applied directly.

In this work, the design is co-optimized in algorithm and hardware architecture to reduce the complexity and achieve high throughput. Bilinear quarter pixel approximation, together with a directional 5T12S (5 transform points among 12 search points) search pattern is proposed to reduce the complexity of the interpolation and search process. Furthermore, an exhaustive size-hadamard transform (ES-HAD) is introduced to improve coding quality, and determine the best transform size rather than using complex transform coding. Besides, a data reuse method of ES-HAD is applied to reduce the hardware overhead. This design is implemented in 65nm CMOS chip and verified by FPGA based evaluation system. It achieves 995Mpixels/s for 7680x4320 30fps encoding, at least 4.7 times faster than previous designs. Its power dissipation is 198.6mW at 188MHz, with 0.2nJ/pixel power efficiency. Despite high complexity in HEVC, the chip achieves 56% and 90% improvement on power efficiency than previous works in H.264 (Tsung,ICME'2009, Kao,TVLSI'2010).

Chapter 5 [Conclusion] concludes the proposed works and present the contribution of this dissertation.