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 Low Complexity Optimization for Video Encoding

申 請 者

ZHANG, Tianruo

情報生産システム工学専攻
マルチメディアシステム研究

2011 年 5 月

Digital video is widely used and plays an important role in the ambient systems of modern society. Many international video coding standards are developed in order to compress the huge data of video signals to make it suitable to be stored and transmitted in TV broadcasting, video conferencing, videophone, video entertainment and surveillance systems. H.264/AVC is the latest video coding standard which provides high compression efficiency while introducing heavy computation load, which leads to high power consumption in real applications.

The high computation complexity is one of the main challenges of developing H.264/AVC video encoder for high definition TV (HDTV) or portable devices. Rate-distortion optimization (RDO) based intra and inter prediction consumes most of the computational complexity of H.264/AVC encoder. Inter prediction consumes most of the computation load. Intra prediction also introduces heavy computation complexity, while its high data dependency causes throughput problem in hardware real-time HDTV H.264/AVC encoder.

The target of this dissertation is to reduce the computational complexity of HDTV 1080p H.264/AVC encoder while keeping the compression efficiency, so as to achieve high performance encoding under low energy consumption. Firstly, a fast mode decision algorithm and a fast block type decision algorithm for intra prediction are proposed for computational complexity reduction as well as throughput enhancement to satisfy the requirement of HDTV real-time encoding. VLSI architecture is implemented to prove proposed fast mode decision algorithm to be an efficient accelerator for intra encoder chip. Then, the computational complexity of the H.264/AVC encoder is reduced by a region-of-interest (ROI) based scheme. A novel hardware-oriented fast ROI detection algorithm is applied in this ROI based approach. Simulation results show that the ROI based scheme reduces encoding time by 77.18%. This dissertation consists of 6 chapters as follows:

Chapter 1 [Introduction] gives a brief introduction to H.264/AVC video coding. The basic knowledge of intra prediction and ROI based coding is discussed, followed by the motivation

and contribution of this work.

Chapter 2 [A Fast Mode Decision Algorithm of Intra Prediction] presents a fast algorithm and its VLSI implementation for reducing computational complexity and increasing throughput of intra encoder. Firstly, the different methods to reduce the processing time of RDO based mode decision in intra prediction are discussed. Then, a hybrid algorithm which consists of a direction detection method with a novel edge detection pattern and a mode prediction technique using spatial neighboring correlation is proposed. RDO only performs on an average of 2.50 modes among all 9 modes in each 4×4 block. Proposed algorithm can further reduce the encoding time by 24% compared with Pan's fast mode decision algorithm. VLSI architecture of proposed intra mode decision module is designed with TSMC 0.18μm CMOS technology. Pipeline architecture is implemented to reduce the number of necessary clock cycles to process each macroblock (MB), so as to enhance the throughput of intra encoding. High operation frequency, low power, small chip area and efficient processing cycle reduction make this design to be an excellent accelerator for HDTV real-time encoder.

Chapter 3 [A Fast Block Type Decision Algorithm of Intra Prediction] presents a fast algorithm which is complementary to the fast mode decision algorithm in Chapter 2 to reduce the computational complexity of intra prediction in high profile by skipping some of the block types. The block type of intra prediction is related to the smoothness of MB. DCT transform is applied to analyze the frequency feature by its DC and AC coefficients. Proposed algorithm selects only one or two block types among 4×4, 8×8 and 16×16 block types. Simulation results show that the time reduction rate is 1.23 times of Lin's work while the performance degradation is lower. This algorithm shows good performance on different resolution from CIF to HDTV 1080p.

Chapter 4 [A Region-of-Interest Detection Algorithm] presents an ROI detection algorithm and its VLSI architecture to detect faces as ROIs for ROI based video coding. The detection algorithm has three main advantages to satisfy the requirements of a pre-processor in video encoder: low power, high accuracy and

MB order processing. The MBs in ROIs are detected sequentially in the same order of H.264/AVC encoding so that the ROI detector and H.264/AVC encoder can be arranged in MB level pipeline. ROI detection is performed in a novel estimation-and-verification process with a novel ellipse contour template. Proposed algorithm and its VLSI architecture can be configured to detect either single ROI or multiple ROIs in each frame and the hardware detection speed of single detection mode is much higher than multiple detection mode. Proposed contour template is implemented by hardware by very simple circuits so that the hardware cost is only 4.68k NAND gates. The detection accuracy is much higher than previous fast ROI detection algorithm while the hardware cost is much lower than face detection algorithms. Proposed architecture performs as an MB level pre-processor with negligible chip area and power consumption overhead of H.264/AVC encoders for variant resolution, from CIF to HDTV 1080p.

Chapter 5 [Region-of-Interest based Encoding Complexity Reduction Scheme] presents an algorithm to reduce computational complexity by using the ROI detector in Chapter 4. The relation between ROI portion and encoding complexity reduction is discussed at first. By using proposed ROI detection algorithm, 42% H.264/AVC encoding time is further reduced compared with using previous ROI only because of ROI size reduction. In order to further increase the time reduction rate and reduce performance degradation, we propose a complexity reduction scheme which consists of 3 methods: (1) the inter prediction mode selection based on quality difference, (2) unequal performance degradation based on unequal bits allocation, and (3) the ROI boundary enhancement to reduce the coding complexity of ROI. Experimental results show that proposed complexity reduction scheme can reduce 77.18% simulation time based on proposed ROI detection when QP difference between ROI and non-ROI is 20, with 1.02% bit-rate increase and 0.04dB PSNR loss. 30.79% of encoding time is further reduced compared with Liu's work with similar performance degradation. Subjective quality of encoded video contents is also enhanced due to the ROI based unequal bit allocation.

Chapter 6 [Conclusion] concludes this dissertation.