

摘要

纵观数字电视、新一代移动通信、宽带网络通信、家庭消费电子这些蓬勃发展的高新技术产业群，其共性技术集中在以音视频为主要内容的多媒体信息处理技术。对这些海量的多媒体信息，特别是视频数据，进行数字化处理的需求促使了视频数据压缩编码近年来在技术及应用方面都取得了长足进展。近几年国际上新制定的 H.264/AVC 视频编码标准以及国内的 AVS/AVS+ 数字音视频编码标准，与以往任何编码标准相比，其压缩效率提高 2~3 倍，编码效率的提高是以高计算复杂度和存储复杂度为代价的。目前通用处理器已不能满足视频编码实时性的要求，超大规模集成电路成为高性能视频编码首选平台之一，所以面向硬件的高效视频编码算法及芯片结构的设计变得尤为重要。

本论文以高性能视频编码技术为研究对象，包括运动估计和视频编码 SoC 芯片结构两个方面，针对影响运动估计计算复杂度的主要因素及视频编码算法的特性，提出了一系列面向硬件的优化算法与芯片结构，以获得在资源消耗(包括存储带宽、芯片面积)、编码效率与系统性能等多种约束下较为权衡的芯片实现方案。

本文的主要内容包括以下四个方面：

1) 计算复杂度约束条件下的动态搜索窗控制 (CDSR) 算法。为达到降低运动估计计算复杂度、合理分配计算资源的目的，提出了一种计算复杂度约束条件下的动态搜索窗控制算法，包括 PMV 偏差度评价算法、SR 与 MVD 分布模型、运动估计计算模型、计算复杂度线性预测模型，该算法根据当前帧的运动复杂度来分配硬件计算资源，可以在有限的计算资源下最大化实时视频编码的率失真性能。本文提出的 CDSR 算法与其等价的全搜索算法相比能够获得 0.1~0.3dB 的性能提升，与其他动态搜索窗算法 (DSR) 相比可以获得 50%~90% 的计算复杂度降低。

2) 二进制自适应像素截断 (BALM) 算法。为了满足对低复杂度运动估计的需求，基于图像局部像素的相关性，提出了一种二进制自适应像素截断 (BALM) 算法并同时给出其 VLSI 结构实现。提出的 BALM 算法与其它像素截断算法相比具有更高的率失真性能，在 $NTB = 4$ 时，BALM 算法可节省 37.41% ASIC 逻辑门数，且编码性能损失小于 0.1dB。

3) 高性能多分辨率运动估计 (MMEA) 算法与 VLSI 结构。为了降低运动搜索算法的复杂度，提高运动估计算法的处理性能，提出了一种面向硬件的高性能多分辨率运动估计算法与 VLSI 结构。运动估计算法采用三层分辨率、由粗至细逐步细化的搜索策略。VLSI 结构采用采用 Level C+ 数据复用策略、双路弓形扫描 (HnSS)、可重配处理单元阵列实现系统吞吐率、片外带宽、片上面积三者间的平衡。提出的 MMEA 运动

估计模块可在 200Mhz 系统工作频率满足 1080P@60fps、2 参考帧的高清视频实时编码。

4) AVS/AVS+高清视频编码 SoC 芯片与高性能帧级与宏块级混合编码流水线结构。提出了在算法性能、硬件资源等多约束下的 AVS/AVS+高清视频编码 SoC 芯片的整体结构，其中包括合理的系统软硬件分区、高效的帧级与宏块级混合视频编码流水线。该 SoC 结构可在 170Mhz 下满足 4 路 PAL/NTSC 标清、单路 1080i@60 高清 AVS/AVS+编码格式视频的实时编码。

关键词：视频编码，AVS，运动估计，芯片结构，流水线结构

Study on VLSI Efficient Motion Estimation Algorithm and Chip Architecture for Video Coding

Xianghu Ji (Microelectronics and Solid State Electronics)

Directed by Prof. Xiaodong Xie

ABSTRACT

In throughout the vigorous development of digital television, new generation wireless communication, broadband network, home consumer electronics, the generic technology of these high-tech industries focus on the processing of multimedia information with audio/video as the main content. With the demand of digitizing those large scale of multimedia information, especially the video data, the video compression technology and the video application have gone through a rapid development recently. The emerging digital video coding standards, such as H.264/AVC and AVS/AVS+, can achieve 2~3 times coding efficiency compared with all the previous coding standards. But the improvement is achieved at the expense of high computational complexity. At the moment, general purpose processor can not meet the real-time requirement of new standards and very large integrated circuit (VLSI) has become the first choice for the implementation of high performance video encoder. Hence, hardware-oriented coding algorithms and VLSI architectures have become more and more important.

This paper targets on the research of video coding technology, including high performance motion estimation and video coding SoC architecture. In order to obtain a good balance among hardware resource consumption (including memory bandwidth, silicon area etc.), coding efficiency and system throughput requirements, this paper proposed a series of hardware-oriented algorithms and corresponding VLSI architectures based on the characteristic of motion estimation and video coding algorithm.

This paper includes the following work:

- 1) Computation-Constrained Dynamic Search Range (CDSR) algorithm. In order to achieve computational complexity reduction and computing resource allocation for motion estimation, this paper proposed a computation-constrained dynamic search range (CDSR) algorithm which includes PMV error evaluation algorithm, SR and MVD distribution model, motion estimation computation model and linear model

for computational complexity prediction. The CDSR algorithm can allocate the computing resources according to the motion complexity in the current frame and maximize the rate-distortion performance on a hardware platform with limited computing resources. It can achieved about 0.1~ 0.3dB average PSNR improvement when the computation consumption is restricted to a specific level as compared with its equivalent Fixed SR algorithm and can achieve about 50%~90% computation savings when compared to the benchmarks.

- 2) Binary Adaptive Luminance Mapping (BALM) algorithm. This paper proposed a Binary Adaptive Luminance Mapping algorithm by using a dynamic mapping for each block which is based on the local pixel correlation of an image and give an architecture for its VLSI implementation. Experimental results show that our proposed BALM achieves higher rate-distortion (RD) performance compared with Bit Truncation (BT) algorithm and PSNR degradation is relatively small when $NTB \leq 5$ using our scheme. And, the NTB4 BALM can achieve 37.41% silicon area saving and power consumption reduction with just PSNR loss of 0.1 dB in our proposed IME architecture.
- 3) High Performance Multi-Resolution Motion Estimation Algorithm (MMEA) and its VLSI architecture. In order to reduce the complexity for motion estimation, a hardware-oriented high performance Multi-Resolution Motion Estimation Algorithm and its VLSI architecture were proposed in this paper. The proposed MMEA searches for the best integer MVs based on the idea of making an initial estimate at the coarse level (resolution) and refining the estimate at the fine level. The VLSI architecture adopt Level C+ data reuse strategy, Horizontal N Snake Scanning (HnSS), re-configurable processing element (PE) array and achieve a good balance among system throughput, external memory bandwidth and silicon area. As a result, the proposed MMEA VLSI architecture can support real time processing of 1080P@60fps with 2 reference frames at 200Mhz.
- 4) High definition video coding SoC architecture and high performance hybrid frame and macroblock level video coding pipeline for AVS/AVS+ video coding standard. This paper proposed an high definition video coding SoC architecture including optimized software-hardware partition, hybrid video coding pipeline for AVS/AVS+ under the constraints of algorithm performance, hardware resource, system throughput. The proposed SoC architecture can support real time coding of 4 channel PAL/NTSC SD or 1 channel 1080i@60 HD AVS/AVS+ video stream

with at working frequency of 170Mhz.

KEY WORDS: Video Coding, AVS, Bit Truncation, VLSI Architecture, Pipeline Architecture