

一种面向 HEVC 运动估计的高性能 VLSI 架构

陈伟伟, 朱 惠, 何卫锋, 毛志刚

(上海交通大学 微电子学院, 上海 200240)

摘 要: HEVC 是正在开发的下一代视频编码标准. 为了进一步提高视频图像的压缩效率, HEVC 中引入了不对称划分模式, 且最大编码单元达到 64×64 像素. 针对 HEVC 标准, 提出了一种面向超高清视频图像应用的高性能运动估计 VLSI 架构. 该 VLSI 架构能够支持从 4×4 到 64×64 的 849 种块的划分种类, 其中包括 168 种不对称划分, 并可重构支持 3 种搜索区间. 通过层次化数据重用策略和存储访问策略, 与传统设计相比将片外数据带宽减少了 99.2%. 采用 TSMC 90 nm 工艺综合, 该电路的最高工作频率可达到 330 MHz, 电路规模 86 万门, 功耗 229 mW. 仿真结果表明, 在 280.4 MHz 工作频率下, 能够完成超高清视频图像 3840×2160 p, 30 f/s 的实时编码要求.

关键词: HEVC; 视频编码; 运动估计; VLSI

中图分类号: TN402

文献标识码: A

文章编号: 1000-7180(2013)06-0010-05

High Performance VLSI Architecture for HEVC Motion Estimation

CHEN Wei-wei, ZHU Hui, HE Wei-feng, MAO Zhi-gang

(School of Microelectronics, Shanghai Jiao Tong University, Shanghai 200240, China)

Abstract: High Efficiency Video Coding (HEVC) is the next generation video coding system under standardization. HEVC employs the Largest Coding Unit (LCU) up to 64×64 pixels with Asymmetric Motion Partitions (AMP), which significantly increases the algorithm complexity of Motion Estimation (ME). In this paper, a novel VLSI architecture compatible to HEVC ME algorithm is proposed for the first time. A quarter of LCU size's Processing Element (PE) array architecture is adopted to conduct full search ME with 849 different block sizes from 4×4 to 64×64 pixels, 168 AMPs such as 4×16 and 24×32 pixels, as well as 3 search ranges (SR) including $[-4, 4]$, $[-8, 8]$ and $[-16, 16]$ pixels. Moreover, a hierarchy data reuse scheme and a memory access strategy are adopted to reduce off-chip bandwidth up to 99.2% compared to the traditional method. Using TSMC 90 nm 1P9M technology, the proposed architecture is synthesized at the maximum work frequency of about 330 MHz with 860×10^3 Gates and 229 mW. Simulation results show that the architecture is able to process 3840×2160 p video at 30 f/s at nearly 280.4 MHz.

Key words: HEVC; video coding; motion estimation; VLSI

1 引言

随着数字视频服务在图像精度与质量上的要求越来越高, 能够适应更大屏幕、有更高的分辨率和更广泛的色域的超高清电视技术标准现已问世^[1]. 超高清图像的数据量巨大, 为了降低超高清电视巨大的数据传输带宽, ISO/IEC 动态图像专家组 (MPEG) 与 ITU-T 视频编码专家组 (VCEG) 提出了最新的视频编码标准^[2-3]: 高效率视频编码标准

HEVC. HEVC 的核心目标是在 H. 264 的基础上将压缩效率提高一倍, 即在保证相同的视频图像质量前提下, 视频流的码率减少 50%.

与现有的视频编码标准相比, 如 H. 264, 其宏块大小最大为 16×16 像素, 而 HEVC 第一次引入了最大尺寸为 64×64 像素的最大编码单元 (LCU) 的概念, 并且其编码单元 (CU) 是从 4×4 到 LCU 的可变的各种尺寸. 除此之外, HEVC 还首次提出了不对称划分 (AMP) 模式, 从而引入了长宽比为 1:4/

3:4 及 4:1/4:3 的不对称块。

运动估计是视频编码过程中运算量最为密集的部分,并且其结果对视频编码的压缩效率以及编码性能有巨大的影响。研究表明,近年来面向运动估计的 VLSI 架构设计在多方面取得了很大进展并得到了实际应用。

为支持不同应用及不同的搜索区域(SR),文献[4]中提出了一种能够支持不同的 SR 的可重构的可变块运动估计 VLSI 架构。为降低搜索区域的输入数据带宽,文献[5]中提出了一种基于 2 层数据复用策略的运动估计架构,能够完成多参考帧、固定块大小的运动估计^[6]。文献[7]中基于现有的数据复用,提出了针对 H. 264/AVC 的运动估计架构。此外,文献[8]中提出了一种能够达到 100% 硬件利用率的针对 H. 264/AVC 可变块运动估计架构。当前多数的运动估计 VLSI 架构是面向 H. 264/AVC 视频编码标准的可变块运动估计架构。然而,针对需要更大的块尺寸、更多的块划分模式、可变的搜索区域、附加的不对称划分模式以及针对 UHDTV 应用的运动估计,对 VLSI 架构的设计提出了更多的挑战。

本文首次提出了一种面向 HEVC 的运动估计 VLSI 架构,该架构支持 849 种可变块的划分,从 4×4 像素到 64×64 像素,包括 168 种 AMP,例如 4×16 像素及 32×24 像素的块,并可重构的支持 $[-4,4)$ 、 $[-8,8)$ 、 $[-16,16)$ 三种搜索区域。

2 HEVC 运动估计

在 H. 264/AVC 视频编码标准中,以 16×16 像素大小的宏块为基本编码宏块;而在 HEVC 视频编码标准中,基本编码单元为大小可变的 CU。HEVC 中 CU 的种类及 AMP 模式如图 1 所示,其中 N 的范围为 $[4,64]$ 。

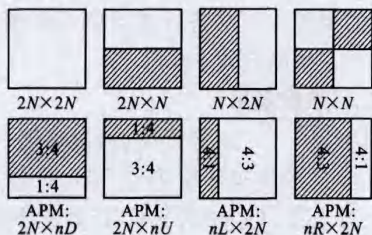


图 1 HEVC 中 CU 划分及 AMP 模式

图 1 中上半部分为 AMP 模式关闭时的划分模式,下半部分为 AMP 模式开启时增加的划分模式。CU 的最大尺寸与 LCU 相同,最小尺寸由划分深度

(Partition Depth, PD) 决定, PD 可变范围为 $0 \sim 4$; 当 LCU 为 64×64 像素, PD 为 4 时, N 的值包括 4、8、16、32, 当 PD 为 0 时,不对 LCU 进行划分。此外 AMP 模式下还包括宽长比为 1:4、3:4、4:1、4:3 的 CU。因此,当 LCU 为 64×64 像素 PD 为 4 时, HEVC 中共有 849 种不同的 CU 划分模式,其中包括 168 种 AMP 划分。在 HEVC 帧间预测时,对上述各种尺寸的块逐一进行 SAD 运算,并通过计算所得的各参考块的 SAD 值及其运动向量偏移位置计算得到 RD Cost, 根据 RD Cost 的最优值的比较选取得到最终的编码划分模式。

3 VLSI 架构及数据复用方案

3.1 运动估计 VLSI 架构

为了能够支持 HEVC 视频编码标准中运动估计的从 4×4 到 64×64 像素的对称及不对称块,并支持可变的搜索区域,本文提出了一种高性能的运动估计架构,如图 2 所示。

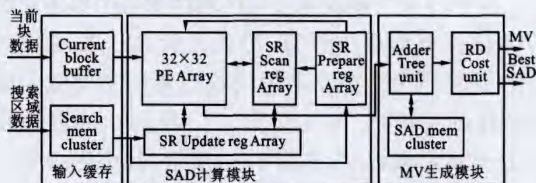


图 2 HEVC 运动估计的 VLSI 架构

该架构由三个主要模块构成:输入缓存, SAD 计算模块及运动向量生成模块。其中输入缓存模块包括用来缓存当前块输入数据的 current block buffer, 用来缓存 SR 数据的 SR mem cluster 以减少从片外取数据的带宽。SR mem cluster 由 6 个单端口 SRAM 组成,为本文的第二层数据复用提供硬件支持。SAD 计算模块由一个 32×32 的 PE 阵列、一个 32×3 的 SR scan reg 阵列、一个 1×35 的 SR update 阵列及一个 32×4 的 SR Prepare reg 阵列组成。上述运算及寄存器阵列合作的设计利用文中即将介绍的第一层的数据复用,并保证了 SR 数据传输与运算的流水性。MV 生成模块中,加法树及 RD cost 计算模块完成 SAD 值的累加和决定最匹配的参考块的 MV; SAD mem cluster 为边长超过 32 像素的块缓存其子块的 SAD 值,用以得到大块的 SAD 值。

3.2 SAD 计算模块与第一层数据复用

图 3 中展示了 SAD 计算模块具体结构及第一层数据复用。同一个当前块的两个左右相邻或上下相邻的 $N \times N$ 像素参考块有 $(N-1) \times N$ 像素的搜

索区域是重叠的,如图 3(a)所示,对这 $(N-1) \times N$ 像素重叠数据的复用为第一层数据复用。

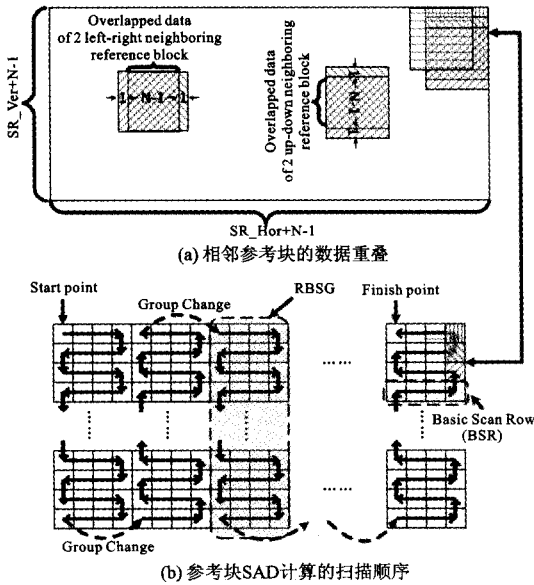


图 3 第一层数据复用

由于一个当前块与搜索区域内的各个参考块的 SAD 值计算为逐个计算,因此设计各个参考块的 SAD 计算顺序是提高参考块数据复用率及减少硬件消耗的关键点. 本文提出了一种新颖的参考块 SAD 计算扫描顺序,如图 3(b)所示,其中图中每一个方格代表一个 $N \times N$ 像素的参考块,所有的方格组成了所有参考块的在搜索区域的映射图. 为数据复用率及硬件面积考虑,将搜索区域映射图划分为宽度为 4 个参考块,高度为 $(2p-1)$ 的参考块扫描组(Reference Block Scan Group, RBSG). 图 3(b)中的箭头展示了参考块之间的 SAD 扫描计算顺序,其中第奇数组 RBSG 扫描顺序为自顶向下的蛇形逐行扫描,第偶数组 RBSG 扫描顺序为自底向上的蛇形逐行扫描。

在上述的扫描顺序下,不同的参考块的数据流水地进入 PE 阵列进行 SAD 运算从而同时利用到了左右相邻以及上下相邻的参考块的数据复用. 其数据复用率为

$$\text{数据复用率} = \frac{(SR_Ver-1) \times \frac{N-1}{N} + \frac{N-1}{2N}}{SR_Ve}$$

其中 SR_Ver 为搜索区域的纵向高度,能够可重构的设置为 8, 16 或者 32. N 为参考块的长度或宽度,式中为 32. 由上式可知,当搜索区域为 $[-16, 16]$ 时, SR_Ver 为 32 参考数据的复用率能够达到 96.6%。

根据上述的参考块 SAD 计算扫描策略,本文提出了针对该策略的 VLSI 架构,该架构由一个 $32 \times$

32 的 PE 阵列及相关的三组缓存阵列构成,如图 4(a)所示. 其中,由 32×32 个运算单元构成的 PE 阵列负责生成当前块与参考块的每一个像素的绝对值差; SR Update Reg 阵列及 SR Scan Reg 阵列缓存参考块数据,用来保证在一组 RBSG 中进行 32×32 的块的运动估计时避免巨大的数据传输带宽; SR Prepare Reg 阵列缓存参考块数据,用来保证在 RBSG 切换时的流水性和数据复用。

SR Scan Reg 阵列由从 $sReg(00, 00)$ 到 $sReg(31, 02)$ 的 32×3 个寄存器构成,该阵列协同 PE 阵列一起动态的缓存和传输搜索区域的参考块数据以进行 SAD 计算. 在每一基本扫描行(Basic Scan Row, BSR)中, PE 与 SR Scan Reg 之间的参考块数据逐一传输;在 RBSG 切换时, PE 与 SR Scan Reg 之间的参考块数据向左侧跨越 4 个 PE/Reg 进行传输。

SR Update reg 阵列为由从 $uReg00$ 到 $uReg34$ 共 35 个寄存器构成的一维线性缓存阵列,负责为下一个 BSR 更新搜索区域的参考块数据;所用时钟周期数与 BSR 扫描所需时钟周期数相同,为 4 个时钟周期. 介于上述的扫描策略,左右扫描顺序在每一个 BSR 之后交替变换,上下扫描顺序在每一个 RBSG 之后交替变换,对应的 SR Update reg 阵列为 PE 阵列的数据更新与上述变换相对应。

SR Prepare Reg 阵列由从 $pReg(00, 00)$ 到 $pReg(31, 03)$ 的 32×4 个寄存器构成. 该寄存器阵列在每一组 RBSG 进行 SAD 计算时,提前为下一组 RBSG 的 SAD 计算提前准备 32×4 像素的参考块数据. 当 RBSG 切换时,该阵列中的数据在一个时钟周期内提供给 PE 阵列及 SR Scan Reg 阵列进行 ME。

通过上述结构,所提出的 VLSI 架构在进行 ME 运算时,能够有效地利用上下、左右相邻的参考块之间的重叠数据,并且 BRS 及 RBSG 的切换均不会引起因需要等待未准备好的数据而停顿,从而使电路工作在较高有效率。

3.3 第二层数据复用

当 LCU 为 64×64 像素时,通过将该块划分为 4 个 32×32 的子块来依次进行 SAD 计算. 图 5 展示了各个 32×32 的子块的 SAD 计算顺序及搜索区域数据的第二层的数据复用。

将 64×64 的 LCU 划分为 32×32 的子块后,每个子块都根据 3.2 节中的数据扫描顺序进行 SAD 运算,并通过子块的 SAD 值累加得到最终的 64×64 的 SAD 值. 从图 5(a)可以看到,每个 64×64 LCU

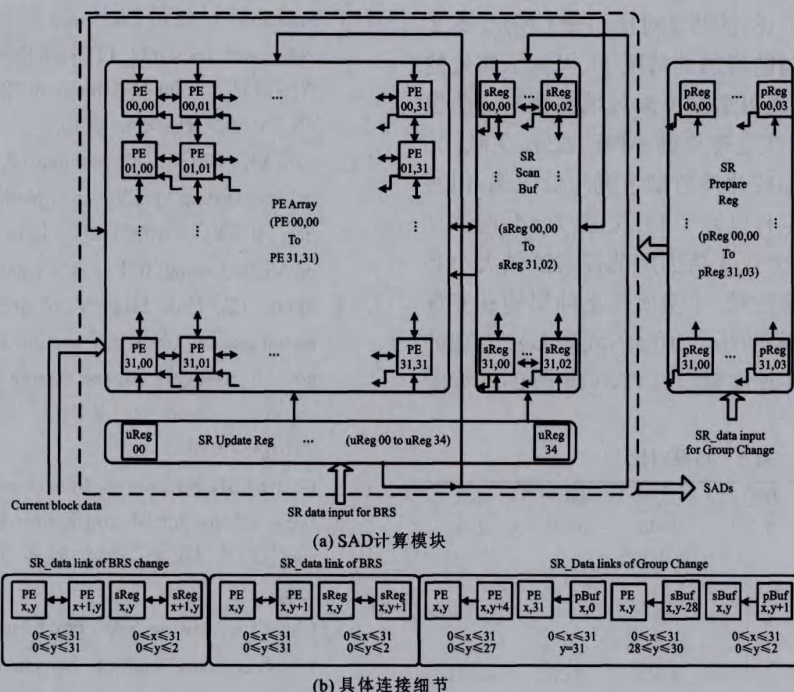


图 4 SAD 计算模块

中的 4 个子块可以数据复用的重叠部分. 其中, 每两个左右或上下相邻的子块的搜索区域有 $(SR_Hor-1) \times (SR_Ver+N-1)$ 或 $(SR_Ver-1) \times (SR_Hor+N-1)$ 像素的数据复用, 而一个 LCU 中四个子块的搜索区域有 $(SR_Ver-1) \times (SR_Hor-1)$ 像素的数据复用; 其中 SR_Hor 及 SR_Ver 分别为搜索区域的纵向高度及水平宽度, N 为子块的长、宽, 为 32. 通过利用上述数据重叠部分从而达到降低片外数据带宽的目的, 因此将 LCU 中的四个子块 SAD 计算顺序设置为 1、2、4、3. 此外, 64×64 像素 LCU 的 SAD 计算扫描顺序设置为列扫描顺序.

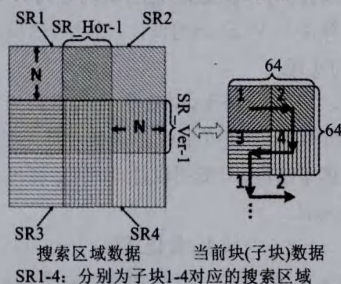


图 5 第二层数据复用及当前块的扫描顺序

为利用上述各个 32×32 的当前块之间数据重叠部分, 并降低运动估计片外数据传输带宽, 本架构采用了搜索区域存储群 (SR mem Cluster) 来缓存搜索区域数据. 通过访存策略的优化, 该存储群仅需要

6 个 1 kB 的端口 SRAM 来达到上述的数据复用. 在 ME 过程中, 4 个 SRAM 用以读出数据来进行 SAD 运算, 2 个 SRAM 写入数据用来为下一个子块的 SAD 运算准备数据; 每一次当前子块的切换, 都至少有 2 个 SRAM 的数据是与上一子块复用的.

上述扫描策略所达到的数据复用率为

$$1 - \frac{4 \times N \times N}{(\text{Frame Height} + SR_Ver - 1) \times (2 \times N + SR_Hor - 1)}$$

其中, 本文所设计电路能够课重构的支持 SR_Ver 为 8、16 或 32, N 为 32. 当搜索区域为 $[-16, 16]$ 时, 即 SR_Ver 为 32 时, 针对 3840×2160 p 的超高清视频图像而言, 数据复用率能够达到 98.03%. 从而与传统的运动估计电路设计相比, 将片外数据带宽降低了 99.2%.

4 电路实现与结果比较

对上述 VLSI 结构进行实现后, 利用 HEVC 标准 C 模型对仿真结果进行验证. 利用 Synopsys Design Compiler 工具, TSMC 90 nm 工艺标准库进行综合, 该结构最高工作频率能够达到 330 MHz, 电路规模为 860 kGates. 对 HEVC 标准码流进行测试处理, 如果考虑 3840×2160 p, 30 f/s 的视频实时

编码,则只需工作在 280.4 MHz 的工作频率下。

与文献[4,5,7,8]的结果对比如表 1 所示。本文在合理的面积开销的增加了的情况下,得到了高效的性能,能够实时处理可重构搜索区域的超高清图像的运动估计,并首次支持块的 AMP 划分。文献[5]和本文架构有相似的片外数据带宽降低比例,但所支持的块尺寸固定且仅为 4×4。文献[7-8]能够支持可变的块尺寸,但仅支持固定的搜索区域大小和普通高清视频图像处理。本文所提出的架构在实现 HEVC 视频编码标准的运动估计运算后,在合理的硬件开销下达到了比文献[4,5,7,8]更好的视频图像编码处理能力。

表 1 结果对比

比较项	文献[4]	文献[5]	文献[7]	文献[8]	本文
PE 阵列	16×16	4×4	16×16	16×16	32×32
搜索区域	[0]— [−8,8]	[−9,9]	[−32,32] × [−16,16]	[−8,8]	[−4,4] [−8,8] [−16,16]
块尺寸	4×4 - 16×16	4×4	4×4 - 16×16	4×4 - 16×16	4×4 - 64×64
AMP 模式	无	无	无	无	有
工作频率	149.3	265.2	216	55.6	330
kGates	N. A.	127.83	168	176	860
工艺库/μm	0.13	N. A.	0.18	0.18	0.9
带宽降低	N. A.	95.7%	N. A.	N. A.	99.2%
片上 RAM/ kB	N. A.	3.96	23.75	41.6	14
处理能力	N. A.	1 280×720 @56 f/s	720×576 @30 f/s	1 280×720 @60 f/s	3 840×2 160p @30 f/s

N. A.:无数据

5 结束语

本文首次提出了一种新型的针对最新视频编码标准 HEVC 的高性能整像素运动估计架构。通过搜索区域数据复用策略及 SAD 值的复用策略,所提出的架构能够支持 4×4 像素到 64×64 像素的块的运动估计,支持不对称的 AMP 划分模式,并能够可重构的支持三种不同的搜索区间:[−4,4],[−8,8]及 [−16,16]。基于上述架构,能够满足在最新视频编码标准 HEVC 下超高清视频应用的实时处理需求。

参考文献:

[1] ITU. Parameter values for ultra-high definition television systems for production and international pro-

gramme exchange. 2012.
[2] Sullivan G J, Ohm J R. Video Coding of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11 (JCTVC - A200)[C]//Joint Collaborative Team on Video Coding (JCT-VC). Dresden, 2010.
[3] Kim I K, McCann K, Sugimoto K, et al. High efficiency video coding (HEVC) test model 7 encoder Description (JCTVC-I1002)[C]//Joint Collaborative Team on Video Coding(JCT-VC). Geneva, 2012.
[4] Peng Li, Hua Tang. VLSI architecture design for reconfigurable block size motion Estimation[C]// Digest of technical papers international conference on consumer electronics (ICCE), 2010. Las Vegas, NV, 2010:439-440.
[5] Grellert M, Sampaio F, Julio C, et al. A multilevel data reuse scheme for Motion Estimation and its VLSI design[C]// ISCAS, 2011. Riode Janeiro, 2011: 583 - 586.
[6] Chen T C, Huang Y W, Tsai Chuan-Yung, et al. Single reference frame multiple current macroblocks scheme for multi-frame motion estimation in H. 264/AVC [C]// ISCAS, 2005. Kobe, 2005:1790-1793.
[7] Zheng Zhaoqing, Sang Hongshi, Huang Weifeng, et al. High data reuse VLSI architecture for H. 264 motion estimation[C]// ICCT, 2006. Cuilin, 2006:1-4.
[8] Li D X, Zheng W. Architecture design for H. 264/AVC integer motion estimation with minimum memory bandwidth[J]. IEEE TCE, 2007, 53(3):1053-1060.

作者简介:

陈伟伟 女,(1988-)硕士。研究方向为 HEVC 视频编码中的运动估计算法及全搜索运动估计算法高性能 VLSI 架构设计、验证与实现。
朱 惠 男,(1989-)硕士研究生,研究方向为 HEVC 视频编码系统的算法与 VLSI 结构设计,以及大规模集成电路的设计、测试与实现。
何卫锋 男,(1976-)博士,副教授。研究方向为视频信号处理的 VLSI 结构、低功耗可重构多核处理器结构、计算机系统结构与微体系结构、IP 复用设计、IP 质量度量与评测、SoC 设计与系统集成。
毛志刚 男,(1963-)博士,教授。研究方向为超大规模集成电路设计方法学、IC 卡芯片设计和硬件安全技术研究、集成电路可测性设计和容错设计技术研究、可编程硬件、可重构硬件结构设计技术。