

中图法分类号: TP368.2 文献标志码: A 文章编号: 1006-8961(2011)05-0740-06

论文索引信息: 罗军, 黄启俊, 常胜, 李昌盛. H.264 中整数变换与量化的 FPGA 实现 [J]. 中国图象图形学报, 2011, 16(5): 740-745

H. 264 中整数变换与量化的 FPGA 实现

罗军, 黄启俊, 常胜, 李昌盛

(武汉大学物理科学与技术学院, 武汉 430072)

摘要: H.264 以其优异的压缩比率和高图像质量在实时网络视频通信、数字广播电视及高清视频存储播放等方面获得广泛应用。变换量化作为 H.264 编码框架中的一个基础模块,是熵编码前的一个重要处理过程,它的主要作用是使输入系数间数据相关性降低。鉴于之前大部分的变换量化是基于软件或协处理器来实现以及此种实现方式在速度及吞吐量上的局限,而硬件实现在速度和吞吐量上则具有很大的优势,因此研究 H.264 变换量化的硬件实现具有实用价值。采用高速并行处理的架构,基于寄存器传输级(RTL)用硬件描述语言完成了 H.264 中的整数离散余弦变换(IDCT)及量化算法的实现,并用 Altera 公司的 Cyclone II 系列可编程逻辑器件实现了硬件验证测试。设计方案消耗了 10 489 个逻辑单元,最高工作时钟频率为 184.88 MHz,数据处理能力达到 2 958 Mpixels/s,可在一个时钟周期之内完成对一个 4×4 矩阵数据的变换量化处理,可满足高速高吞吐量数据流处理的要求。

关键词: 整数离散余弦变换; 量化; H.264 标准; 可编程逻辑器件(FPGA)

A FPGA implementation of integer discrete cosine transform and quantification for H. 264 compression

Luo Jun, Huang Qijun, Chang Sheng, Li Changsheng

(School of Physics and Technology, Wuhan University, Wuhan 430072 China)

Abstract: Due to its excellent compression rate and high image quality, H.264 is widely applied in many regions, such as real-time internet video communication, digital television broadcast, HDTV and so on. As a key component of the H.264 coding framework, integer discrete cosine transform (IDCT) and quantification module takes on the rule of reducing the correlation coefficients of input data, which is an important pre-processing before the entropy coding. The normal realizations are based on software or hardware co-processor, in which way speed and throughput are limited. As a contrast, realization based on hardware can make up these shortages. In this paper, an FPGA implementation of integer discrete cosine transform and quantification for H.264 compression is introduced, which is based on a high throughput structure and is accomplished by register-transfer level (RTL) description. The whole design is tested on the Cyclone II-based development board, DE II. The resource cost is 10 489 LEs. Its maximal work frequency is 184.88 MHz, the maximal throughput is 2 958 Mpixels/s, and it can process a 4×4 input data within a single clock period.

Keywords: IDCT; quantification; H.264 compression standard; FPGA

0 引言

H.264 以其优异的性能在实时网络视频通信、

数字广播电视及高清视频存储播放等方面获得广泛应用,其目标是基于高视频分辨率,提高图像质量,并能够覆盖所有低带宽和高带宽的应用。H.264 视频编码标准编码框架中主要包括以下几个部分:预

收稿日期:2010-02-05;修回日期:2010-04-30

基金项目:国家科技支撑计划(2008BAC36B05)。

第一作者简介:罗军(1986—),男。武汉大学物理科学与技术学院微电子与固体电子学 2009 级硕士研究生,主要研究方向为微电子 CAD 技术。E-mail:kyye168@126.com。

测编码,变换量化和反变换反量化,环路滤波以及熵编码。变换量化的主要作用是使输入系数间数据相关性降低,是熵编码前的一个重要处理过程。H.264采用的变换是类似离散余弦变换(DCT)的无乘法整数变换,量化是采用0~51级的分级量化。变换编码可以去除原始图像的空间冗余,使图像能量集中在一小部分系数上,这样不仅可以提高压缩比而且还可以增强抗干扰能力。H.264中变化量化具有精度高,可通过加减移位来实现乘法以及消除编解码的不匹配问题等优点。可编程逻辑器件(FPGA)具有可编程的灵活配置性,拥有丰富的寄存器资源以及一次性费用低等特点,相比ASIC实现的一次性费用高以及基于DSP处理数据吞吐量低而言,基于FPGA来实现变换量化的算法具有明显的优势。目前国外的研究中,Ihab Amer等人^[1]就给出了一种关于H.264整数变换与量化的硬件结构,并在赛灵思的Virtex II器件中进行了验证,在共消耗了29 018个LUT的资源下,得到了14.598 ns的关键延时,不过时钟频率只有68.5 MHz。Bruguera和Osorio^[2]给出了一种转置矩阵的方式来实现变换量化,采用AMS 0.35 μm 工艺进行综合后得到最高67 MHz的处理速度。除此之外还有基于软硬件的协同设计的方式等。台湾地区Heng等人^[3]针对H.264中的变换量化及反变换量化提出了一种低复杂度的硬件架构,采用TSMC 0.35 μm 工艺进行了综合,结果显示在编码阶段可以以32 MHz的速度处理256 M/s的数据。刘海鹰^[4]介绍了一种基于FPGA的H.264变换量化的硬件原型的设计,以Atera公司的stratix II系列的EP2S15器件为目标综合后的最高频率为147.02 MHz。Zhang等人^[5]介绍了一种组合直接型的2维变换和量化结构,采用Altera公司的Cyclone II器件作为目标器件,并且用Synplifypro进行了综合,处理一个 4×4 的矩阵需要20个时钟周期,共消耗了1 289个组合逻辑单元,可以到达最大85.1 MHz的时钟频率。何云壮等人^[6]介绍了一种采用FPGA实现整数DCT变换模块的设计方案,基于Altera公司的Cyclone系列器件可在4个时钟周期内完成一个 4×4 块的2维整数DCT。陈瑛等人^[7]给出了一种包含正反变换量化硬件结构,在ISE的XST下综合后最大综合频率可达112 MHz,处理一组 4×4 的残差数据的变化量化需要45个时钟周期。Peng等人^[8]基于ASIC方式给出了一种并行处理块架构,以2维的变换量化为基本处理

引擎,可用于流水线加速处理,能够达到156 M/s的数据吞吐量。另外,Wang等人^[9]给出了一种包含反变换反量化的基于Virtex 4 Pro FPGA的硬件架构,并且分别给出了基于面积和速度的优化结果,处理速度大于100 MHz。上述文献在大容量,高速度及高带宽的应用领域中,变换量化硬件实现的处理性能仍有有待改进的地方。文献[1]虽然给出了不同架构的变换量化的硬件实现,但处理速度都没有超过100 MHz。文献[2-3]都是基于ASIC的方式来实现的。文献[4]基于时分复用的技术对一个 4×4 的矩阵进行变换量化仍需要多个时钟周期。文献[5]中处理一个 4×4 的矩阵需要20个时钟周期,而频率只有85.1 MHz。文献[6]只完成了变换的功能。文献[7]中虽达到了高于100 MHz的处理速度,但仍需要45个时钟周期处理一个 4×4 的矩阵。针对上述问题,提出一种新的H.264中整数变化与量化模块的FPGA实现,采用高速并行的流水线处理技术,提升了此模块的性能,实现了处理速度要高于150 MHz并可在一个时钟周期之内完成一个 4×4 的矩阵的处理。

1 系统设计思路

H.264中整数变换及量化,如果输入块是色度块或帧内 16×16 预测模式的亮块,则将宏块中各 4×4 块的整数余弦变换的直流分量组合起来再进行哈达码变换,进一步压缩码率。图1给出了H.264对输入矩阵数据的处理过程。Intra_16 \times 16 luma表示 16×16 的亮度数据,8 \times 8 chroma表示8 \times 8的色度信号。输入 4×4 的数据残差跟随残差块的标记信号线一起输入到变换量化模块,首先进行前向整

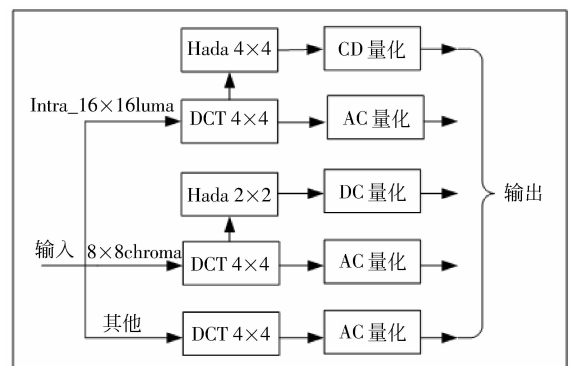


图1 变换量化的数据流程图

Fig. 1 Data flow diagram of transform and quantization

数离散余弦变换,再对变换后的结果进行判定是否进行哈达码变换;对前向整数离散余弦变换和哈达码变换后的数据分别进行交流系数量化和直流系数量化。由变换量化的处理流程,硬件设计的思路是把整个模块划分为三大部分:整数 DCT 变换,哈达码变换以及量化。为了提高处理速度实现对每个时钟周期处理一个 4×4 的矩阵,设计中对于每一个输入的数据采用流水线来处理,同时在整体实现上对 16 个数据采用并行处理。

2 系统设计实现

2.1 系统框图

整个系统模块划分为三大部分:整数离散余弦变换,哈达码变换及量化。系统模块划分及数据流框图如图 2 所示。框图中对所有的输入矩阵先进行前向整数 DCT 变换,若输入矩阵为亮度块的 16×16 的帧内模式或色度块,则需要对前向整数 DCT 变换后的系数再进行哈达码变换,变换后的系数经过量化之后得到输出结果。设计划分为 7 个模块,其中前向整数离散余弦变换 (IDCT) 模块完成前向整数 DCT 变换;系数提取 (EXE) 模块的主要功能是完成对 IDCT 模块后的系数进行提取,以便进行哈达码变换; 4×4 Hada 模块完成 4×4 的哈达码变换功能; 2×2 Hada 模块完成 2×2 的哈达码变换功能;AC_Quant 模块的主要功能是完成前向整数 DCT 变换后系数的量化;DC_4_Quant 模块和 DC_2_Quant 模块分别完成 4×4 哈达码变换和 2×2 哈达码变换后系数的量化。

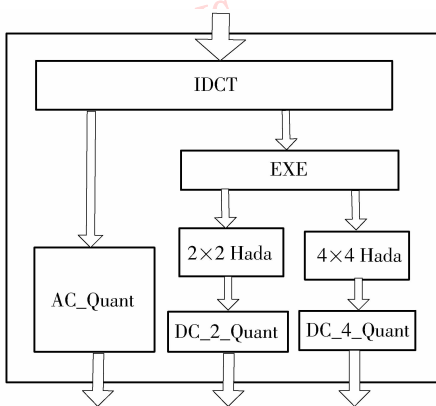


图 2 系统模块划分图

Fig. 2 Partition map of system module

2.2 主要模块实现及综合结果

2.2.1 前向整数 DCT 变换模块

设计中对于整数 DCT 变换采用并行和流水线的处理方法,如图 3 所示。采用蝶形算法分别对输入的数据进行行和列的变换,这样经过 5 个时钟周期后可以得到变换后的输出,并且每个时钟周期可以处理一个矩阵的数据。这样的设计可以大大提高处理数据的速度。图中的每列或每行都是采用相同的蝶形算法来实现,共有 5 级流水线。

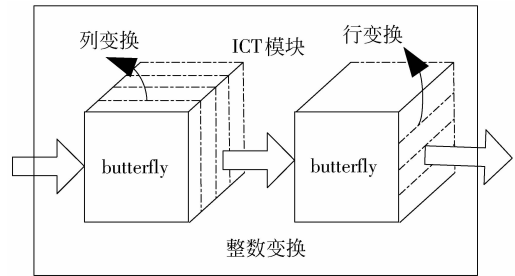


图 3 整数 DCT 变换原理图

Fig. 3 Integer DCT transform diagram

1 维列变换蝶形算法如图 4 所示。列变换和行变换的蝶形算法的实现是基本相同的,只是输入数据的顺序以及输入位宽不同。第 1 个蝶形算法模块的输入位宽为 9 位;第 2 个蝶形算法模块的输入位宽 16 位。 4×4 Hada 模块和 2×2 Hada 模块基于流水线的处理架构,采用相似的蝶形算法来实现,不再详述。

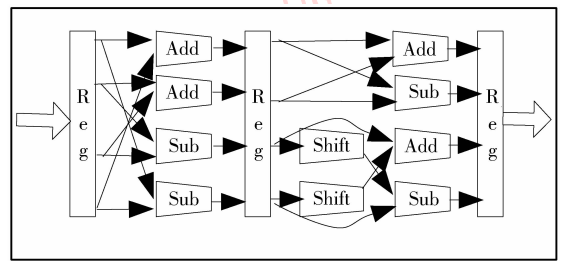


图 4 蝶形算法的流水线硬件实现

Fig. 4 Pipelined hardware implementation of butterfly algorithm

2.2.2 交流系数量化模块

设计中把每个数据的量化过程划分为 3 个大的部分:系数生成,相乘和相加移位。而对于共 16 个系数采取并行处理的方式,即相当于模块复用。量化的数据流框图如图 5 所示。系数生成模块中主要采用了查找表和 ROM 存储系数常量,乘法模块采用 Quartus II 自带的 IP 核。数据处理流程采用了流

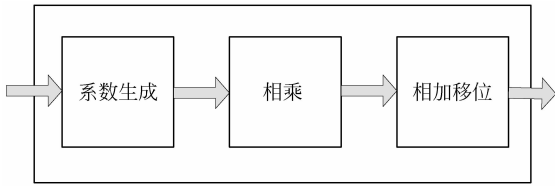


图 5 量化过程框图

Fig. 5 Diagram of quantification

流水线的方式以便提高处理速度。更为详细的模块处理过程如图 6 所示。

量化过程采用了 5 级流水线,使用寄存器来保存中间结果。输入的系数根据其其在 4×4 矩阵中的位置来选择不同的相加相乘的系数常量进行乘、加或移位操作。Qf_0 为输入变换后的数据, row_0 和 col_0 分别为行列指示信号, QP 为量化参数, Intra 为模式指示信号, Q_result 为量化后的输出

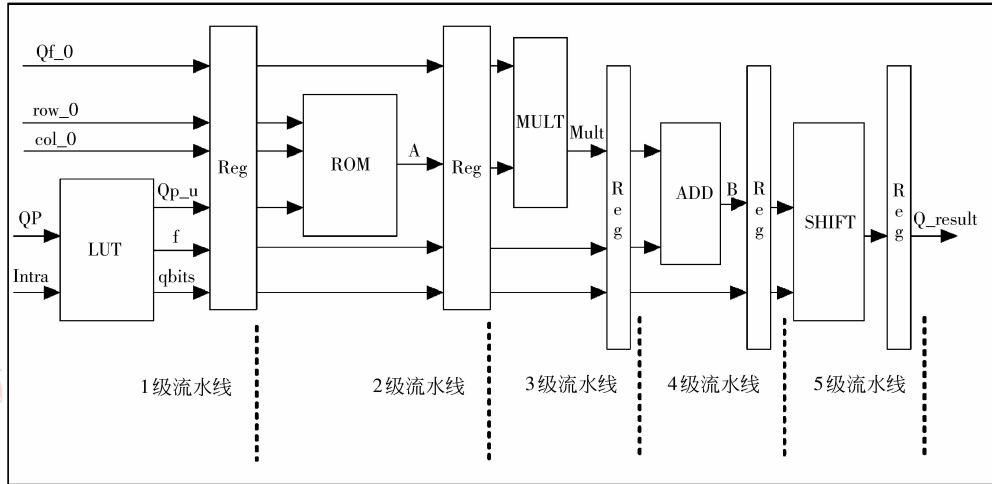


图 6 流水线实现量化结构图

Fig. 6 Pipeline structure diagram of quantification

数据。而由于 4×4 的 DC 系数量化过程和 2×2 的 DC 系数量化过程与 AC 系数的量化处理过程类似,因此也采用了与交流系数量化处理模块相同的硬件架构来实现,在具体实现的过程中用查找表代替 ROM 模块来产生系数 A。

2.2.3 系统综合

整个变换量化系统以 Cyclone II 系列器件为目标器件在 Quartus II 7.1 下进行了综合。由于处理的数据比较多,引脚消耗比较大(在 H. 264 编码框架中,设计的输入输出端口信号都是作为内部信号使用),因此在实际验证时量化后的输出系数用 8 位位宽来表示。综合后的资源使用情况如表 1 所示。表 1 中列出了系统的变换量化和仅包含前向整数 DCT 变换交流量化的综合结果。系统在约束了输入输出端口信号的位宽后,以 Stratix II 系列 EP2S15 为目标器件进行了综合,表 2 列举了此综合后结果。相比文献[4]中的设计,此设计在速度及吞吐量上都有了较大的提升。表 3 中列举了本文和文献[6]中整数 DCT 变换的综合结果。

表 1 设计的综合结果

Tab. 1 Result of the project synthesize

模块名称	器件型号	资源使用情况	最高工作频率
系统变换量化	Cyclone II	逻辑单元 10 489 (15%)	184.88 MHz
		存储器 28 672 bits(6%)	
前向 DCT 变换量化	EP2C35F672C6	逻辑单元 2 497 (10%)	196.12 MHz
		存储器 28 672 bits(6%)	

表 2 设计结果比较

Tab. 2 Comparison of design results

器件	资源消耗 (ALUTs)	最高工作 频率/MHz	处理一个 4×4 块时间
文献[4]	—	147.02	多个时钟周期
本文	5 949	269.25	1 个时钟周期

表 3 整数 DCT 综合结果

Tab. 3 Result of integer DCT synthesize

器件	逻辑 单元	存储器	最高工作 频率/MHz	处理一个 4×4 块时间
文献[6]	EP1C6Q240C8	1 632	256	—
本文	EP2C35F672C6	1 242	0	271.08

钟频率为 184.88 MHz,可在一个时钟周期之内处理一个 4×4 的图像矩阵,消耗逻辑资源为 10 489 个。设计可以满足高速高吞吐量处理的要求。

参考文献 (References)

- [1] Amer I, Badawy W, Jullien G. A proposed hardware reference model for spatial transformation and quantization in H.264 [J]. Journal of Visual Communication and Image Representation, 2006, 17(2):533-552.
- [2] Bruguera J D, Osorio R R. A unified architecture for H.264 multiple block-size DCT with fast and low cost quantization [C]//Proceedings of the 9th EUROMICRO Conference on Digital System Design (DSD'06). New York: Inst. of Elec. and Elec. Eng. Computer Society, 2006:407-414.
- [3] Heng Yaolin, Yi Chihchao, Che Hongchen, et al. Combined 2-D transform and quantization architectures for H.264 video coders [C]//Proceedings of the International Symposium on Circuits and Systems. New York: Institute of Electrical and Electronics Engineers Inc., 2005:1802-1805.
- [4] Liu Haiying, Zhang Zhaoyang, Shen Liquan. A high-performance hardware implementation of the H.264 transformation and quantization oriented to FPGA [J]. Journal of Image and Graphics, 2006, 11(11):1636-1639. [刘海鹰,张兆杨,沈礼权.基于 FPGA 的 H.264 变换量化的高性能的硬件实现[J].中国图象图形学报,2006,11(11):1636-1639.]
- [5] Zhang Qidong, Li Ji, Cao Xixin, et al. A novel algorithm and architecture of combined direct 2-D transform and quantization for H.264 [J]. The Journal of China Universities of Posts and Telecommunications, 2007, 14(Sup):79-83.
- [6] He Yunzhuang, Liu Yongqiang, Li Yongquan. Implementation of integer DCT in H.264 on FPGA [J]. Microcomputer Information, 2007, 17(6):205-206. [何云壮,刘永强,李勇权.H.264 整数 DCT 的 FPGA 实现[J].微计算机信息,2007,17(6):205-206.]
- [7] Chen Ying, Zhao Gang, Su Haibing. High performance architecture for transform and quantization of H.264 based on FPGA [J]. Modern Electronics Technique, 2009, 32(10):19-21. [陈瑛,赵刚,苏海冰.一种基于 FPGA 高性能 H.264 变换量化结构设计[J].现代电子技术,2009,32(10):19-21.]
- [8] Peng Chungan, Yu Dunshan, Cao Xixin, et al. A new high throughput VLSI architecture for H.264 transform and quantization [C]//Proceedings of the 7th International Conference on ASIC. New York: Inst. of Elec. and Elec. Eng. Computer Society, 2007:950-953.
- [9] Wang Leirui, Zhang Zhaoyang, Teng Guowei, et al. Hardware implementation of transform and quantization for AVS encoder [C]//Proceedings of the International Conference on Audio, Language and Image Processing. New York: Inst. of Elec. and Elec. Eng. Computer Society, 2008:843-847.