

H.264 帧内 4×4 块预测模式选择及其 IP 核设计

吴从中, 李本斋, 胡有刚

(合肥工业大学 计算机与信息学院, 安徽 合肥 230009)

摘要: 提出了一种基于抽样的 H.264 帧内 4×4 块预测模式选择方法, 设计了基于 PLB 总线的 IP 核, 对预测点进行抽样选择, 并对硬件资源的消耗进行了优化。将此方法在 JM 中与全搜索进行了比较, 并将设计的软核在 Xilinx Virtex-II PRO 开发板中进行了验证。结果表明, 该抽样算法在 PSNR 和编码比特数方面都能达到与全搜索相近的结果, 而且软核在硬件资源的消耗上也有明显的减少, 编码效率明显提高, 能够适应实时编码。

关键词: H.264; 帧内预测模式选择; 抽样; PLB; 软核

中图分类号: TP302

文献标识码: B

Intra-prediction mode selection for 4×4 blocks in H.264 and IP design

WU Cong Zhong, LI Ben Zhai, HU You Gang

(School of Computer Science and Information Engineering, Hefei University of Technology, Hefei 230009, China)

Abstract: This paper advanced a method for intra-prediction mode selection for 4×4 blocks in H.264 based on sampling and designed its IP based on PLB bus, which are optimized for the selection prediction pixel and hardware resources. This method has been compared with full search in JM, and the IP is verified on a Virtex-II PRO FPGA. The results show that it can receive similar result compared with FS in PSNR and bits and consumed very little hardware resource. With high efficiency, this design could fully adapt to real-time coding.

Key words: H.264; intra-prediction mode select; sample; PLB; IP

新一代视频编码标准 H.264 以其高复杂度为代价获得了优异的编码效率, 但运算复杂度的增加给实时编码带来了很大困难, 因此 H.264 编码器的硬件化是必然趋势。随着 FPGA 技术的发展, 用 SoPC 实现 H.264 视频信息的实时编码是一种极佳的选择。本文介绍的 H.264 帧内 4×4 块预测模式选择及其 IP 核是基于 SoPC 的 H.264 编码器的一个硬件加速器。

帧内预测是 H.264/AVC 的一个重要组成部分, 它充分利用了帧内图像的空间相关性, 提高了压缩效率, 对编码器整体性能的提高具有重要作用。在帧内预测编码过程中, 预测块 P 基于已编码重建块和当前块形成。对亮度像素而言, P 块用于 4×4 子块或者 16×16 宏块的相关操作。其中, 4×4 亮度子块有 9 种可选预测模式, 独立预测每一个 4×4 亮度子块, 适用于带有大量细节的图像编码。16×16 亮度块有 4 种预测模式, 预测整个 16×16 亮度块, 适用于平坦区域图像编码; 色度块也有 4 种

预测模式, 类似 16×16 亮度块预测模式。编码器通常选择使 P 块和编码块之间差异最小的预测模式。实验表明, H.264 的帧内预测约占编码总时间的 29%, 提高了帧内预测的速度, 对实现实时编码具有重要的意义。

本文结合实际应用, 通过对帧内预测模式选择原理的分析, 根据图像空间方向的相关性, 提出了一种适合于硬件实现的算法, 并将其设计成 IP 核, 作为 PowerPC 的一个硬件加速模块在 Xilinx Virtex-II PRO 开发板中进行了验证。

1 基于抽样的帧内预测模式选择方法及硬件结构

FS (全搜索) 能达到很好的预测效果, 而且易于硬件实现, 但计算量过大。本文提出的基于抽样的模式选择算法, 不仅易于硬件实现而且减少了计算量。

1.1 帧内预测的原理^[1]

如图 1(a) 所示, 4×4 亮度块包括标示为 $a\sim p$ 的 16 个待预测像素, 与之相邻的左上方标示为 $A\sim M$ 为已编码

硬件纵横 Hardware Technique

并重构的像素,帧内预测就是利用 $A \sim M$ 来预测 $a \sim p$ 。根据预测方向的不同,共有 9 种可选预测模式,其中 DC 预测(模式 2)根据 $A \sim M$ 中已编码像素预测,即当 $A、B、C、D、E、F、G、H、I、J、K、L$ 都可用时,用 $(A+B+C+D+I+J+K+L) \gg 3$ 来表示其所有像素;若 $A、B、C、D$ 可用而 $I、J、K、L$ 不可用,用 $(A+B+C+D) \gg 2$ 来表示,反之用 $(I+J+K+L) \gg 2$ 表示,若都不可用,则预测值为 128。其余 8 种模式的预测方向如图 1(b)中的箭头表示。模式 0 是垂直预测,即 $a=e=i=m=A$;模式 1 为水平预测,即 $a=b=c=d=I$;模式 3~8 的预测像素由 $A \sim M$ 不同加权平均得到。

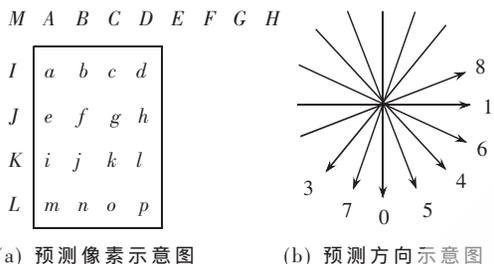


图 1 4x4 亮度预测

确定最佳预测模式的代价函数模型有 RDO 和 SAD 两种,其中 RDO 计算方法比较复杂且不易于硬件实现。本文使用 SAD 计算方法,其计算主要是加减平均等运算,适合于硬件实现。其计算公式为:

$$Cost_{16} = SAD_{16} + 4R\lambda(Qp) \quad (1)$$

其中, $4R$ 是对使用某种预测模式后比特数的估计; $\lambda(Qp) = 0.85 \times 2^{(Qp-12)/3}$, 当 Qp 一定时为常数; SAD (绝对差和) 是 16 个像素预测值与图像像素值的差值。

1.2 基于抽样的帧内预测模式选择

在 H.264 帧内预测模式选择中, 4×4 块每种模式判别时需要计算 16 个像素点的预测值,还要计算 16 个像素点的 SAD 值,计算量较大。同时,如果一次输出 16 个预测值,则预测电路消耗的硬件资源很大,且在 H.264 的后续处理中需要大容量的存储器与之匹配。如果一次只输出一个像素,消耗的资源会减少,但对硬件的频率要求变高。在一个时钟周期预测一个像素情况下,如果要实现 SDTV (720x480x30 fps) 的实时处理,频率就要达到 202 MHz (720x480x1.5x30x13)。所以,如果对 4×4 块中 16 个像素进行抽样计算,不仅大大减少了计算量,而且对硬件的频率要求也会降低。由于 4×4 块与 16×16 块相比较小,块内纹理变化比较平缓,抽样的像素可以大致代表块内纹理方向。本文将 4×4 宏块抽样为 4 个部分^[2],图 1(a)中 4 部分像素抽样为(1) a, c, i, k ; (2) b, d, j, l ; (3) e, g, m, o ; (4) f, h, n, p 。在这 4 个部分中选取一个部分进行 SAD 失真计算,从而省略了大部分的运算代价。每一组抽样的代价表示为:

$$Cost_4 = SAD_4 + R\lambda(Qp) \quad (2)$$

其中, SAD_4 是每个抽样组的 4 个像素点与图像像素值

的绝对差值和,其他参数同式(1)。

通过以上抽样,可以减少 3/4 的运算量,要实现 SDTV 所需的硬件频率仅为 50.54 MHz (720x480x1.5x30x13/4)。为了验证抽样对图像质量的影响,将这种抽样的方法在 JM 中与 FS 进行了比较。

1.3 硬件结构设计

本文的抽样算法在 FPGA 中实现的系统结构框图如图 2 所示,包括每种模式的预测值计算和 4 点的 SAD 值计算模块。由于 $\lambda(Qp)$ 的计算比较复杂,不易于硬件实现,这里先将计算结果存放在 FPGA 中,通过 Qp 查表的方法获得。

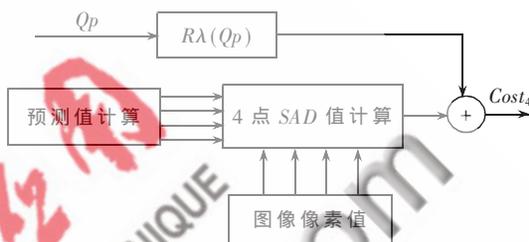


图 2 帧内预测模式选择系统结构框图

H.264 帧内 4×4 块预测模式共有 9 种,每一种都有对应的预测值计算公式。在硬件设计时要为每一种预测模式设计一个预测值计算模块。对于 mode0 (垂直预测)和 mode1 (水平预测),电路的输出直接等于输入;其他的模式计算比较相似,都是通过加法和移位来完成的^[3],这里以 mode3 (上下对角预测)为例给出其硬件结构如图 3 所示。该结构通过对已知像素值进行累加和移位来得到相应的预测像素的值。系统中 9 种预测模式并行计算,使每个计算模块都处在工作状态^[4],提高了运算效率。

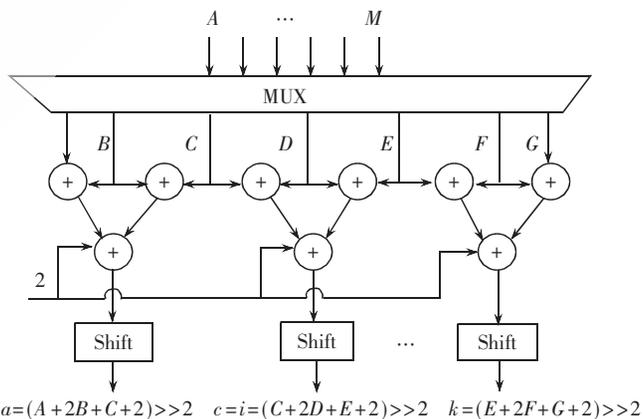


图 3 mode3 抽样像素预测值计算硬件结构

在帧内预测模式选择中, SAD 值的计算最复杂。本文对像素点的抽样只需要进行 4 点的 SAD 值计算即可,降低了计算复杂度和硬件资源的消耗。4 点的 SAD 值计算式如式(3)所示,其硬件结构如图 4 所示。

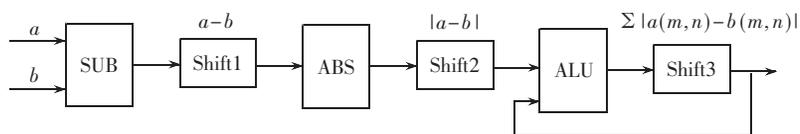


图4 SAD值计算硬件结构

$$SAD_4 = \sum_{m=1}^4 \sum_{n=1}^4 |a(m,n) - b(m,n)| \quad (3)$$

其中, $a(m,n)$ 表示预测值, $b(m,n)$ 为图像像素值。

2 基于 PLB 总线的 IP 核设计与测试

2.1 PLB 总线简介

PLB 总线 (Processor Local Bus) 是 IBM 开发的一种高性能片上总线, 主要应用于 PowerPC405 处理器系统中, 它支持 32 位、64 位和 128 位数据宽度, 本文对 IP 核的设计使用了 64 位总线宽度。虽然 PLB 接口总线协议非常复杂, 但是 Xilinx 为用户专门设计了一套可以用工具生成的接口协议, 称作 plbv46_slave_single。它在用户 IP 核和 PLB V4.6 总线标准之间提供了一个双向的接口。plbv46_slave_single 让用户可以便捷地在 IBM PLB Bus 和用户 IP 核之间进行交互。这个 slave 服务使得用户能在不同地址范围内提供地址译码, 从而配置多个用户 IP 的接口到 PLB Bus 上。plbv46_slave_single 可以优化以减少点到点连接设计的延时和 FPGA 资源消耗。

2.2 基于 PLB 总线的软核设计

H.264 帧内预测模式选择软核的设计包括总线接口和帧内预测模式选择模块两部分。PLB 总线接口使用 Xilinx 为用户设计的 plbv46_slave_single。而帧内预测模式选择模块与硬件体系无关, 方便移植。

H.264 帧内 4×4 块预测模式的系统硬件结构如图 5 所示。输入寄存器由 4 个 32 bit 的寄存器 0、1、2 和 3 组成, 输入已编码并重构的像素 $A \sim M$, 寄存器 3 的高位用于输入 Q_p 值; 寄存器 4 为原始图像抽样出用于 SAD 计算的 4 点像素值; 寄存器 5、6、7、8 为输出数据寄存器; 寄存器 9 为控制寄存器, 包括启动、完成、复位等。预测值计算模块就是根据 $A \sim M$ 对抽样像素进行每种模式的预测值计算, SAD 值计算模块计算每种模式预测值与图像像素值的差值, 第三个模块根据每种预测模式的计算结果选择最优的预测模式输出。其中, $R\lambda(Q_p)$ 根据 Q_p 值

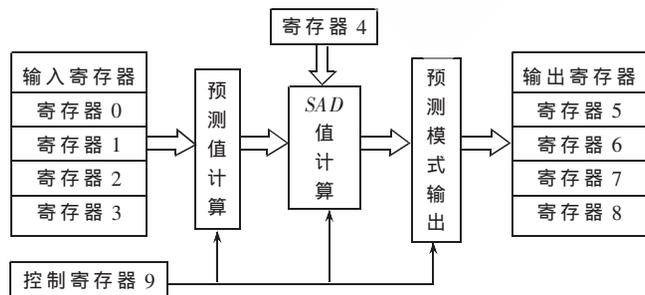


图5 帧内预测模式选择硬件结构

在 FPGA 中查表得到。

2.3 软核的仿真与测试

测试中首先使用 IBM 的 CoreConnect 工具和 Modelsim6.0a 来仿真设计软核, 仿真使用虚拟平台来测试, 仿真的目的是保证 PLB 总线接口能被处理器正确访问。通过虚拟处理器将虚拟内存中的数据写入待测软核, 然后读取待测软核中的数据并判断是否正确。

仿真通过以后, 再将软核集成到系统中, 以便验证软核在实际系统中的工作是否符合要求。软核验证使用 PC 机和目标板相结合的方法, 目标板使用 Xilinx 公司 XUP Virtex-II PRO 开发板, 内部含有 2 个 PowerPC 内核。

验证时, 首先使用 EDK9.1 将软核集成到 PowerPC 系统中, 编译通过后由 USB 接口下载到目标板, 系统开始运行后通过 RS232 从上位机下载待编码的图像数据, 下载的数据保存在目标板上的 256 MB DDR SDRAM 中。数据下载完毕后程序将待编码数据依次写入软核并启动, 将计算完的数据写入 DDR SDRAM, 待全部数据处理完毕, PowerPC 将处理结果一起发送给上位机。上位机将结果与本机 C 代码执行结果相比较, 最终确认软核是否正确工作。

3 综合性能分析

首先在 JM 中对基于抽样的帧内预测模式选择方法进行验证。通过对 Foreman 和 Akiyo 两个视频序列选择不同的 Q_p 值与全搜索进行比较, 结果如表 1 所示。可以看出, 通过这样的抽样并没有带来 PSNR (信噪比) 的明显下降和编码比特数的增加。

表 1 本方法与全搜索比较

序列	Q_p	方法	PSNR/dB	编码比特数/kbit
Foreman 序列	5	全搜索	57.30	430 264
		抽样搜索	57.28	430 300
	48	全搜索	26.61	20 856
		抽样搜索	26.55	20 856
Akiyo 序列	5	全搜索	56.14	39 912
		抽样搜索	56.12	40 292
	48	全搜索	28.06	701
		抽样搜索	27.99	711

同时将该模块在 Xilinx 的 XC2VP30 FPGA 中进行综合, XC2VP30 含有 13 696 个 Slices 和 136 个 18×18 乘法器, 综合工具使用 Xilinx 的 ISE9.1。综合的结果和资源消耗情况如表 2 所示。可以看出, 本文中所设计的模块在关键路径和资源消耗上都优于参考文献 [5] 中所述的方法。

本文将 H.264 的帧内 4×4 块的预测模式选择方法

进行优化。表 1 和表 2 所示的实验数据表明,在不损失图像质量的同时明显地降低了运算的复杂度。同时,与硬件相结合,设计出基于 PLB 总线的 IP 软核,并在 Xilinx XUP Virtex-II PRO 开发板中做了验证。

表 2 软核综合结果

关键路径/ns	5.2
时钟频率/MHz	192.30
Slices	3 409
实现 SDTV(720×480 ×30 fps)所需频率/MHz	50.54

参考文献

- [1] 毕厚杰. 新一代视频压缩编码标准——H.264/AVC[M]. 北京:人民邮电出版社,2005.
- [2] MENG Bo Jun, OSCAR C A. Fast intra-prediction mode selection for 4×4 blocks in H.264. IEEE International Conference on Acoustics Speech and Signal Processing, 2003(3):389-392.
- [3] 黄凯,秦兴,严晓浪,等. 一种 H.264 帧内预测模式判决算法及 VLSI 实现体系[J]. 电子学报, 2007,35(2):207-211.
- [4] HUANG Yu Wen, HSIEH Bing Yu, CHEN T G, et al. Analysis, fast algorithm, and VLSI architecture design for H.264/AVC intra frame coder. IEEE Transactions on Circuits and Systems for Video Technology, 2005, 15(3): 378-401.
- [5] 朱总平,冯建华,曹喜信. H.264/AVC 帧内预测器的 VLSI 实现[J]. 北京大学学报, 2008(1):44-48.

(收稿日期:2009-10-16)

作者简介:

吴从中,男,1965年生,硕士,副教授,主要研究方向:数字信号处理、嵌入式系统技术。

李本斋,男,1984年生,硕士研究生,主要研究方向:视频编码、集成电路设计。

胡有刚,男,1984年生,硕士研究生,主要研究方向:视频编码、嵌入式系统设计。