

一种支持 H.264 和 AVS 的帧内预测器设计

徐张磊 郑世宝 杨宇红

(上海交通大学电子工程系图像通信与信息处理研究所, 上海 200240)

(上海交通大学上海市数字媒体处理与传输重点实验室, 上海 200240)

摘要 为了使多标准视频解码器中的帧内预测器能够支持 H.264 和 AVS 两种视频标准,在对 H.264 和 AVS 两种标准中的帧内预测计算模式进行分析,并对各模式计算公式之间相似性进行分析的基础之上,提出了一种支持 H.264 和 AVS 两种标准的,可配置的帧内预测值计算硬件架构。该架构由于将大部分预测模式的计算放到一个可配置的计算单元中进行,从而大大减少了芯片资源的浪费。为了提高处理速度,可采用 4 个相同的可配置的计算单元并行计算,一次计算出 4 个像素点的预测值。实验结果表明,该硬件架构在 FPGA 上占用 10 371 个 LUTs,频率可以达到 150MHz。

关键词 H.264 数字音视频编解码标准 帧内预测 硬件架构 可配置

中图分类号: TN919.81 **文献标识码**: A **文章编号**: 1006-8961(2007)10-1735-05

A Design of Intra Predictor Supporting H.264 and AVS

XU Zhang-lei, ZHENG Shi-bao, YANG Yu-hong

(Institute of Image Communication and Information Processing, Department of Electronic Engineering,
Shanghai Jiaotong University, Shanghai 200240)

(Shanghai Key Laboratory of Digital Media Processing and Transmissions, Shanghai Jiaotong University, Shanghai 200240)

Abstract To adapt the intra predictor in multi-standard video coder to H.264 and AVS stand, a configurable hardware architecture design of intra predictor that supports both H.264 and AVS was proposed based on the similarities between the intra prediction algorithms in both standards. This kind of architecture can implement most intra prediction algorithms in one configurable calculate unit and thus can reduce the use of chip area. In order to enhance the speed of process, four parallel configurable generator unit are adopted in our design. This design was synthesized and the result showed that this architecture occupied 10 371 LUTs on FPGA and its frequency can be 150 MHz.

Keywords H.264, AVS, intra prediction, hardware architecture, configurable

1 引言

随着多媒体与网络技术的飞速发展,视频压缩编解码技术也得到了极大的进步。国际上先后制定了一系列视频图像编码标准,目前使用最广泛的是 MPEG2 标准,另外,作为新一代视频编解码技术的代表,H.264/AVC 标准^[1]和数字音视频编解码技术标准(advanced audio video coding standard, AVS)^[2]在编码效率和图像质量上远远超过了以往的标准,受到了人们普遍的关注和重视。面对多标准共存,并且在

短时间内很难统一的局面,视频解码芯片的兼容性就自然显得十分重要。因此,设计和实现能够支持多种视频标准的视频解码芯片具有重要的实用价值。

帧内预测解码作为视频解码中重要的一部分,如何设计一个硬件架构来有效地将其实现,是解码器设计中一个极其重要的问题,国内外有很多学者对此进行了研究,并取得了不少进展^[3-6]。但在可以同时支持多种视频标准的解码器中的帧内预测器的设计上,尚研究较少。

本文针对同时支持 H.264 和 AVS 标准的帧内预测解码器的硬件实现进行讨论,在基于现场可编程

收稿日期:2007-06-01; 改回日期:2007-07-01

第一作者简介:徐张磊(1984~),女,现于上海交通大学图像通信与信息处理研究所攻读硕士学位。研究方向为视频编解码、芯片设计。E-mail: xuzhanglei@sjtu.edu.cn

程门阵列(field programmable gate array, FPGA)验证平台的基础之上,设计并实现了一个可用于解码端的支持多标准、可配置的帧内预测器。

2 帧内预测算法简介

帧内预测是预测编码的一种,其主要目的是去除空间上的冗余性。其压缩原理就是利用图像内部特别是图像平滑部分相邻像素间的相关性,用已编码完成的图像块去预测当前图像块,以达到降低编码码率的目的。

H.264 对亮度分量和色度分量设定了不同的预测方案,分别独立进行预测。亮度分量的帧内预测包含基于 4×4 像素块(简称基于 4×4 块)、 8×8 像素块(简称基于 8×8 块)和基于 16×16 像素块(简称基于 16×16 块)的帧内预测等 3 类预测模式。其中基于 4×4 块和 8×8 块的亮度预测各有 9 种预测方式, 16×16 块的亮度预测有 4 种预测方式。色度分量只有一类预测模式,就是基于宏块大小的预测,而色度宏块大小根据取样格式的不同,有 $8 \times 8(4:2:0)$, $16 \times 16(4:4:4)$, $8 \times 16(4:2:2)$ 3 种。色度预测有 4 种预测方式,它们对不同彩色取样格式下的宏块大小均适用。

AVS 标准中的帧内预测思路与 H.264 基本一致,主要在 H.264 之上进行了简化,但同时也对某些模式的算法进行了改进,以达到更好的预测效果。AVS 的帧内亮度预测和色度预测都是以 8×8 的块为单位,而不受宏块大小、彩色取样格式的限制。其中亮度预测有 5 种预测方式,色度预测有 4 种预测模式。

3 帧内预测器的设计和实现

H.264 和 AVS 两者的编解码过程都是基于 16×16 大小的宏块进行的。在解码端,解码器针对每个宏块分别进行解码。解码流程为:首先根据经过熵解码得到的语法元素确定本宏块的预测模式,然后根据此预测模式对本宏块中的像素进行预测计算,此计算所需要的点就是与本宏块相邻的已解码宏块中的像素点。最后将计算后得到的预测值和经过反变换反量化得到的残差值相加来得到重建值,就得到了本宏块像素解码后的像素值。

3.1 帧内预测器架构设计

根据帧内预测器的解码工作流程,整个帧内预

测器主要由预测模式判断和预测值计算两部分组成。另外,由于预测值计算要用到相邻宏块的像素值,因此需要不断读写存储器。由于频繁读写外部存储器会使得系统速度大大下降,因此本设计使用 FPGA 片上 block ram 来存放预测所需要的像素值。本文设计的帧内预测器的系统结构框图如图 1 所示。

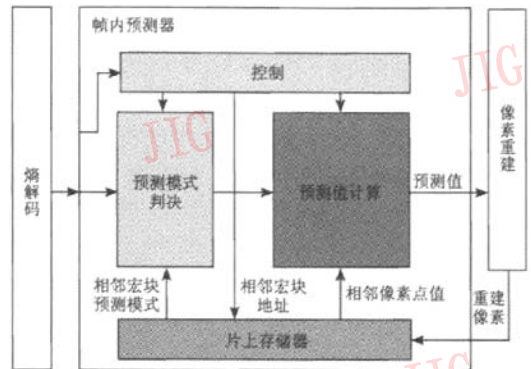


图 1 帧内预测器架构

Fig.1 Architecture of intra predictor

由于在解码端,帧内预测模式已经由编码端决定,因此模式判决并不是帧内预测器的关键模块。而预测值计算则占有重要地位,所以如何快速根据预测模式和相邻像素求出预测值就成为设计解码器的关键。

3.2 帧内预测模式算法分析

通过对 AVS 和 H.264 的帧内预测进行比较可以发现:

在亮度分量的预测方面,无论是块尺寸还是预测算法,AVS 都比 H.264 简化了很多。AVS 没有基于 4×4 块和 16×16 块的预测方式,而只有 5 种基于 8×8 块的预测方式,且基本上属于 H.264 中的 9 种基于 8×8 块预测方式的子集;但是在直流(简称 DC)模式、左下对角线模式这两种模式上的计算方法又和 H.264 标准有所不同。

在色度分量的预测方面,H.264 和 AVS 都具有 4 种基于 8×8 的预测方式,其中除了 DC 模式的计算方法不同外,其他 3 种是一样的。但 H.264 比 AVS 复杂,因为如果彩色取样格式不是 $4:2:0$,H.264 中的色度分量,还有基于 16×16 块或者 8×16 块大小的预测方式,则两标准中所有的预测模式可归纳为 31 种(如表 1 所示)。

表 1 帧内预测方式
Tab.1 Categories of intra prediction

标准	块大小	预测模式	备注
H.264	16 × 16	垂直预测(I16_V)、水平预测(I16_H)、Plane 预测(I16_Plane)	亮度、色度
		DC 预测(I16_DC)	亮度
		DC 预测(I16_DC_chroma)	色度
H.264	4 × 4	垂直预测(I4_V)、水平预测(I4_H)、DC 预测(I4_DC)、右垂直预测(I4_VR)、下水平预测(I4_HD)、下左对角线预测(I4_DDL)、下右对角线预测(I4_DDR)、左垂直预测(I4_VL)、上水平预测(I4_HU)	亮度
	8 × 16	水平预测(I8 × 16_H)、垂直预测(I8 × 16_V)、Plane 预测(I8 × 16_Plane)、DC 预测(I8 × 16_DC)	色度
H.264 和 AVS	8 × 8	垂直预测(I8_V)、水平预测(I8_H)	H.264 亮度、色度
		DC 预测(I8_DC)、下左对角线预测(I8_DDL)、下水平预测(I8_HD)、上水平预测(I8_HU)、右垂直预测(I8_VR)、左垂直预测(I8_VL)	H.264 亮度
		下左对角线(I8_DL)	AVS 亮度
		下右对角线预测(I8_DDR)	H.264 亮度, AVS 亮度
		DC 预测(I8_DC_chroma)	H.264 色度
		Plane 预测(I8_Plane) DC 预测(I8_DC_AVIS)	H.264 色度、AVS 色度 AVS 亮度、色度

如果对每种不同的预测模式都设计相应的预测计算器,那么对硬件资源是极大的浪费。其主要原因是在解码端,由于预测模式是已经确定的,因此几种预测模式不会同时进行。如果对每种模式都设计预测计算器,那么在某种预测模式下,只有一个预测模式在进行计算,而其他的计算器则都处于空闲状态。实际上,由于各种预测模式的计算方法之间有着很大的相似性,因此可以设计一个支持大部分预测模式的可重构的计算单元。

预测计算模式可以分为以下 5 类:(1)直接将邻近像素的值作为预测值,主要在垂直模式、水平模式,以及下水平方向预测的某些情况中出现;(2)基于 $a+2 \times b+c$ 和 $a+b$ 的计算结构,大部分方向预测都是如此;(3)基于 $(a+2 \times b+c) + (d+2 \times e+f)$ 的计算结构,只有 AVS 中的下左对角线模式和 DC 模式需要使用到这种计算结构;(4)DC 模式,需要利用所有邻近像素的值进行计算,但是 H.264 和 AVS 中对 DC 的处理不同,H.264 中所有像素点的 DC 值均相同,而 AVS 中则需要对每个点都进行计算;(5)Plane 模式,需要首先用邻近像素计算出 a, b, c ,再使用 a, b, c 对每个像素点进行计算。其计算比较复杂,下面将予以专门分析。

Plane 模式是一种比较特殊的模式,其计算方式和其他预测模式均不相同。但是根据参考文献[3]中对 Plane 模式所做的分析,则可以将 Plane 模式下复杂的计算方法转化为和其他大部分模式类似的计

算方法。以 16×16 块的 Plane 模式为例进行说明,其计算公式如下:

$$\text{pred}[x, y] = \text{Clip1}((a + b \times (x - 7) + c \times (y - 7) + 16) \gg 5)$$

其中, x, y 表示像素点在宏块中的坐标,取值范围均为 $0 \sim 15$; a, b, c 是通过邻近像素点所计算出的变量值。宏块中第 1 行的预测值计算公式如下:

$$\begin{aligned} \text{令 } A_0 &= a + b \times (-7) + c \times (-7) + 16, \text{ 则} \\ \text{pred}[0,0] &= \text{Clip1}((a + b \times (-7) + c \times (-7) + 16) \gg 5) \\ &= \text{Clip1}(A_0 \gg 5) \\ \text{pred}[1,0] &= \text{Clip1}((a + b \times (-6) + c \times (-7) + 16) \gg 5) \\ &= \text{Clip1}(A_0 + b \gg 5) \\ \text{pred}[2,0] &= \text{Clip1}((a + b \times (-5) + c \times (-7) + 16) \gg 5) \\ &= \text{Clip1}(A_0 + 2 \times b \gg 5) \\ \text{pred}[3,0] &= \text{Clip1}((a + b \times (-4) + c \times (-7) + 16) \gg 5) \\ &= \text{Clip1}(A_0 + 3 \times b \gg 5) \end{aligned}$$

其余像素点的计算公式和上述公式类似。通过对计算公式的分析可以发现,复杂的计算公式可以简化为 $A + i \times b$ 的计算(其中 i 可以取 $0, 1, 2, 3$),即 $A, A+b, A+b+b, A+b+b+b$ 等 4 种;然后对和移位,Clip,即可完成,这种计算方法和大部分预测模式下的方法相同。

3.3 预测值计算模块的设计

各种预测模式的计算方法虽然不同,但是运算类型相近。根据前文对各种预测模式下的计算公式进行的分类容易发现:

$$\begin{aligned}
 a + 2 \times b + c &= (a + b) + (b + c) \\
 (a + 2 \times b + c) + (d + 2 \times e + f) &= \\
 &= (a + b) + (b + c) + (d + e) + (e + f)
 \end{aligned}$$

由于从中可以找到最基本的计算单元类型,就是 $(a + b)$,所以,可以以 $(a + b)$ 为核心,设计结构为 $(a_1 + a_2) + (b_1 + b_2) + (c_1 + c_2) + (d_1 + d_2)$ 的计算单元(如图 2 所示,图中 Clip 为限幅运算)。这样就可将大部分的预测计算集中到一个可重构的运算单元中完成。根据预测模式以及需要计算的像素点在块中的位置,就可对这个运算单元的输入进行选择,可从所有邻近像素中选择 8 个输入。对于将邻近像素的值直接作为预测值的情况,也就是在前文中第 1 类预测模式的情况下,可以选择邻近像素输入 a_1 ,其他输入均为 0。同时,根据对计算模式的分类,在计算之后,再从计算出的值 $v_0, v_1, v_2, v_3, v_{plane}$ 中选择一类计算结构的结果作为输出。这样就使得两个标准中共 31 种预测模式中的绝大部分可以放在一个可重构的计算单元中完成。

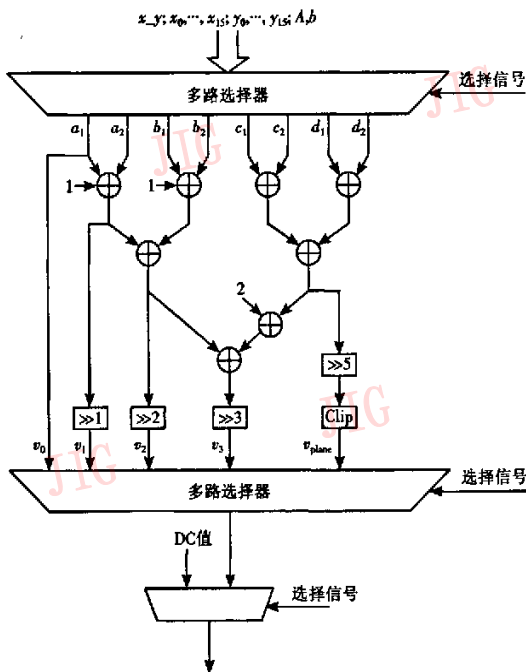


图 2 可重构计算单元

Fig. 2 Configurable calculating unit

通过上文对 Plane 模式的分析可以看出,Plane 模式的计算也可以放到该可重构计算单元中。为了得到计算所需要的中间变量 A 和 b ,可通过设计一个 Plane 预处理模块来先计算出 Plane 模式下所需

要的 a, b, c ,以及每个点所需要的 A 。根据基于块大小的不同,Plane 有基于 8×8 块, 8×16 块以及 16×16 块 3 种,由于其计算公式类型一致,只是其中参数不同,因此,3 种 Plane 模式的预处理可以放到一个模块中完成。

由于 H.264 中 DC 模式需要用到所有的邻近像素点,使用图 2 所示计算单元一次不能计算出来,因此对 DC 模式的计算需要设计单独的 DC 进行模块计算,以便一次将块中所有像素点的 DC 值计算出来。与 H.264 不同,AVS 中 DC 模式的计算可以使用此可重构运算单元进行计算。

为提高处理速度,可采用 4 个同样的可配置计算单元,一个时钟输出 4 个像素点(即图中 pixel)。整体的预测计算单元如图 3 所示。

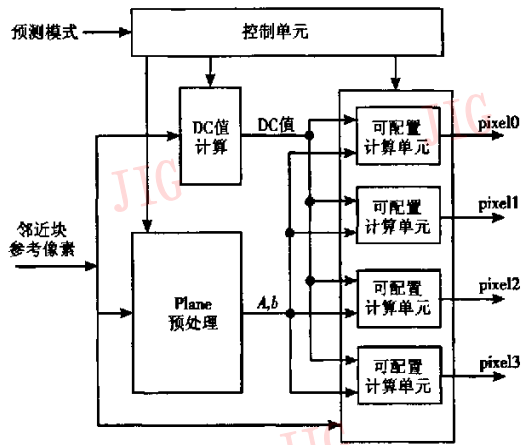


图 3 预测值计算单元

Fig. 3 Prediction calculating unit

4 实验结果

使用 VerilogHDL 语言对本文所提出的预测值计算结构中的预测值计算单元进行实现,并用 ModelSim 对其进行仿真实验,仿真结果如图 4 所示。从图 4 中可以看出,每个时钟输出 4 个像素点的预测值,经过 64h 就可以将一个 16×16 宏块的 256 个像素点的预测计算完成,并输出。

为评估本文的应用效果,选择 Xilinx 公司 Virtex2 系列的 XC2V4000FPGA 芯片,使用 Synplify Pro 软件对本文所提出的预测值计算单元进行综合,频率设置为 150MHz,设定 retiming 和 pipeline 有效。综合后的资源占用情况为:L/O primitives 为

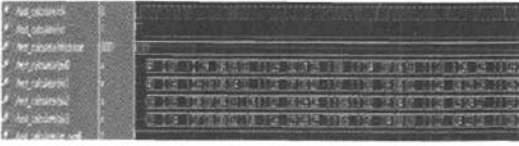


图 4 仿真结果

Fig. 4 Simulation Result

317, Block Multipliers 为 4, Total LUTs 为 10 371。

5 结 论

本文提出了一种支持 H. 264 和 AVS 两种标准的,可配置的帧内预测值计算硬件架构。其主要思想是根据不同预测模式计算公式之间所具有的相似性,将除 DC 模式之外的大部分预测模式的计算都放到一个可配置的计算单元中,以大大减少芯片资源的浪费。为了提高处理速度,本架构采用 4 个相同的可配置的计算单元并行计算,一次可计算出 4 个像素点的预测值。笔者选择 Xilinx 公司 Virtex2 系列的 XC2V4000FPGA 芯片进行综合,频率可以达到 150MHz 以上,可满足高清视频解码的实时要求。本设计不但可以适用于支持多标准的视频解码芯片中,也可以单独用于 H. 264 或 AVS 解码芯片。

参考文献 (References)

- 1 Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG, Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification [S]. ITU-T Rec. H. 264/ISO/IEC 14496-10 AVC, March, 2005.
- 2 GB/T20090.2-2006. Information Technology-Advanced Audio Video Coding Standard Part 2: Video [S]. [GB/T20090.2-2006, 信息技术 先进音视频编码 第二部分:视频 [S].]
- 3 Huang Y W, Hsieh B Y, Chen T C, *et al.* Analysis, fast algorithm, and VLSI architecture design for H. 264/AVC intra frame coder[J]. IEEE Transactions on Circuits and System for Video Technology, 2005, 15(3): 378 ~ 401.
- 4 Liu Ling-zhi, Lu Qi, Rong Meng-tian, *et al.* A parallel design of H. 264 intra predictor generator [J]. Journal of Shanghai Jiaotong University, 2006, 40(1): 54 ~ 58. [刘凌志, 路奇, 戎蒙恬等. 一种并行结构的 H. 264 帧内预测器[J]. 上海交通大学学报, 2006, 40(1): 54 ~ 58.]
- 5 Yang Chen, Li Shu-guo. A high parallel VLSI design of H. 264 intra predictor generator [J]. Microelectronics and Computer, 2006, 23(12): 111 ~ 117. [杨晨, 李树国. 一种高并行度的 H. 264 帧内预测器的 VLSI 设计[J]. 微电子学与计算机, 2006, 23(12): 111 ~ 117.]
- 6 Wang Zheng, Liu Pei-lin. Analysis of AVS intra prediction technology and its implementation by hardware [J]. Computer Engineering and Applications, 2006, 42(19): 80 ~ 83. [王争, 刘佩林. AVS 帧内预测算法及其解码器的硬件实现[J]. 计算机工程与应用, 2006, 42(19): 80 ~ 83.]