

基于 Loeffler 算法的 2-D DCT IP 软核设计

郭宝增, 牛 力, 刘志明

(河北大学 电子信息工程学院, 河北 保定 071002)

摘 要: 提出一种基于 Loeffler 算法的 2-D DCT IP 软核设计方法. 用移位和加法运算代替乘法运算. 为减少芯片占用面积, 对乘法系数采用 CSD 编码, 1-D DCT 复用技术; 为提高电路的速度, 采用流水线结构, 优化转置矩阵. 基于上述算法, 设计了用 Verilog HDL 语言描述的 IP 软核. 对软核进行了编译、综合、布局布线和后仿真, 验证了算法的正确性. 实验结果显示最高工作频率可以达到 139.43 MHz, 能够满足视频图像压缩的实时性要求.

关键词: 1-D DCT; Loeffler 算法; IP 核; 图像压缩

中图分类号: TN47; TP39

文献标识码: A

文章编号: 1000-7180(2011)02-0136-04

Design of 2-D DCT IP Soft Core Based on Loeffler's Algorithm

GUO Bao-zeng, Niu Li, LIU Zhi-ming

(College of Electronic and Information Engineering, Hebei University, Baoding 071002, China)

Abstract: This paper presents a design of 2-D DCT IP soft core based on Loeffler's algorithm. We make use of the shifter and addition operations to replace the multiplication operations. In order to save hardware resources, the multiplication factors are coded by CSD, and the technology of 1-D DCT reusing is taken; In order to improve the circuit speed, the pipelining technology is taken, and optimize the transposed matrices is adopted. Based on above-mentioned algorithms the IP soft core with Verilog HDL is obtained. For the IP soft core, we made the compiling, synthesis, layout and post simulation, so the correctness of the design is verified. The experiment shows that design has a maximum frequency of 139.43 MHz, which can satisfy the requirement of real-time video image compression.

Key words: 1-D DCT; Loeffler algorithm; IP core; image compression

1 引言

离散余弦变换(DCT)是数字图像处理和压缩编码领域最常用的变换之一. 文献[1]提出并实现图像数据从空间域到频率域的变换, 以减少图像数据的相关性, 便于后续的压缩处理. 其性能最接近最佳的 K-L(Karhunen-Loev)变换. 被广泛应用于各种视频压缩标准中. 在图像压缩系统中, 运算量大, 实时性要求高. DCT 的运算量大约占压缩运算总量的 40%左右. 因此设计占有资源少, 实时性高的 2-D DCT IP 核成为一种需要^[2-3].

文中选用 Loeffler^[4]算法. 该算法同其他快速算法相比, 实现八个点的一维 DCT 变换只需要 11

次乘和 29 次加, 乘法器的数量降低到了理论极限. 基于 88 2-D DCT 算法, 设计一个可以在硬件上实现的高速处理视频图像 IP 软核.

2 算法描述

88 点 2-D DCT 变换定义如下:

$$Y(u, v) = \frac{1}{4} c(u) c(v) \sum_{i=0}^7 \sum_{j=0}^7 x(i, j) \cos \frac{\pi(2i+1)u}{16} \cos \frac{\pi(2j+1)v}{16}$$
$$u = 0, 1, \dots, 7; v = 0, 1, \dots, 7 \quad (1)$$
$$c(u), c(v) = \begin{cases} \frac{1}{\sqrt{2}} & u, v = 0 \\ 1 & u, v \neq 0 \end{cases}$$

收稿日期: 2010-01-21; 修回日期: 2010-03-15

基金项目: 河北省自然科学基金(F2009000226)

根据行列分解的思想^[2],二维离散余弦变换硬件实现可以分解为两次一维离散余弦变换,中间加一转置矩阵来实现。因此,提高一维变换的速度是提高二维变换速度的重要基础。

8 点 1-D DCT 变换定义如下:

$$Y(k) = \frac{1}{2}c(k) \sum_{n=0}^7 x(n) \cos \frac{(2n+1)k\pi}{16} \quad (2)$$

$$k = 0, 1, \dots, 7$$

$$c(k) = \begin{cases} 1 & k = 1, 2, \dots, 7 \\ \frac{1}{\sqrt{2}} & k = 0 \end{cases}$$

若直接用定义式(2)计算一维离散余弦变换,需要 64 次乘,56 次加。所需乘法器太多,不能满足实时性要求。如何减少乘法器的数目是减小面积的关键。进一步的研究发现,式(2)的系数矩阵具有特殊的对称性,可以用式(3)来计算 8 点的 1-D DCT 变换,需 22 次乘,32 次加。运算量大幅减少。后来 Loeffler^[4]提出了一种快速一维 DCT 算法,计算 8 点一维 DCT 只需要 11 次乘 29 次加。乘法器的数量降到理论的极限。算法的基本结构如图 1 所示。图 2 为算法中用到的符号,依次为蝶形因子、倍乘因子、旋转因子。

$$\begin{bmatrix} Y_0 \\ Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \\ Y_5 \\ Y_6 \\ Y_7 \end{bmatrix} = \begin{bmatrix} C_0 & C_0 & C_0 & C_0 \\ C_2 & C_6 & -C_6 & -C_2 \\ C_4 & -C_4 & -C_4 & C_4 \\ C_6 & -C_2 & C_2 & -C_6 \\ C_1 & C_3 & C_5 & C_7 \\ C_3 & -C_7 & -C_1 & -C_5 \\ C_5 & -C_1 & C_7 & C_3 \\ C_7 & -C_5 & C_3 & -C_1 \end{bmatrix} \begin{bmatrix} x_0 + x_7 \\ x_1 + x_6 \\ x_2 + x_5 \\ x_3 + x_4 \\ x_0 - x_7 \\ x_1 - x_6 \\ x_2 - x_5 \\ x_3 - x_4 \end{bmatrix} \quad (3)$$

其中 $C_i = \frac{1}{2} \cos\left(\frac{i\pi}{16}\right)$, $i = 1, 2, \dots, 7$ 。

从图 1 可以看到完成一次蝶形运算需要 2 次加法,倍乘因子需要 1 次乘法。旋转因子需要 4 次乘 2 次加,利用式(4)和式(5)对旋转因子做变形,可得式(6)和式(7),可以看到经过变形后实现一次旋转因子只需要 3 次乘 3 次加。

$$y_0 = ax_0 + bx_1 = (b-a)x_1 + a(x_0 + x_1) \quad (4)$$

$$y_1 = -bx_0 + ax_1 = -(a+b)x_0 + a(x_0 + x_1) \quad (5)$$

$$O_0 = \left(K \sin \frac{\pi}{2N} - K \cos \frac{\pi}{2N}\right) I_1 + K \cos \frac{\pi}{2N} (I_0 + I_1) \quad (6)$$

$$O_0 = - \left(K \cos \frac{\pi}{2N} - K \sin \frac{\pi}{2N}\right) I_0 + K \cos \frac{\pi}{2N} (I_0 + I_1) \quad (7)$$

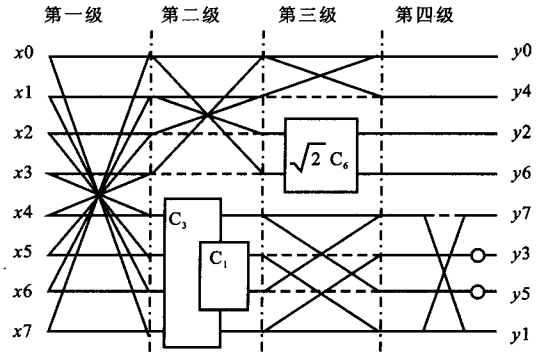


图1 11个乘法器实现8点DCT算法

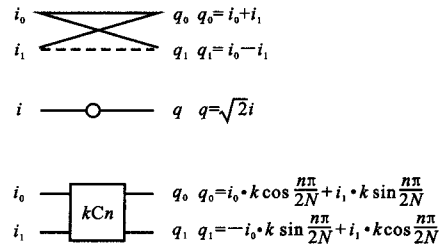


图2 算法中用到的符号

对整个算法做一下统计:

y_0, y_4 是 3 级蝶形因子的级联,共需 6 次加。

y_2, y_6 是 2 级蝶形因子与 1 级旋转因子级联,共需 3 次乘,7 次加。旋转因子 3 次乘可以并行运算,所以最多只需 1 次乘法时间。

y_3, y_5 是 1 级蝶形因子、1 级旋转因子、1 级蝶形因子、1 级倍乘因子的级联,共需 7 次加,5 次乘。旋转因子 3 次乘法并行计算需要 1 次乘法时间,倍乘因子需要 1 次乘法时间。所以需 2 次乘法的时间。

y_1, y_7 是 1 级蝶形因子、1 级旋转因子、2 级蝶形因子的级联,共需 9 次加法,3 次乘。

可知实现偶数部分最多需要 1 次乘法时间,而实现奇数部分需要 2 次的乘法时间。所以对奇数部分进行变换。最终使用图 3 结构进行奇数部分的计算。表 1 为奇数系数值。

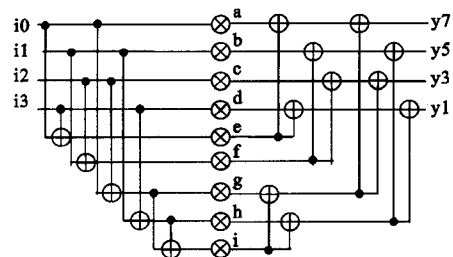


图3 奇数部分的计算

表 1 奇数系数值

奇数系数	值
a	$C_3 + C_5 - C_1 - C_7$
b	$C_1 + C_3 + C_7 - C_5$
c	$C_1 + C_3 + C_5 - C_7$
d	$C_1 + C_3 - C_5 - C_7$
e	$C_7 - C_3$
f	$-C_3 - C_1$
g	$-C_3 - C_5$
h	$C_5 - C_3$
i	C_3

当采用图 3 结构, 奇数部分最多也只需 1 次乘法时间. 实现了奇数部分与偶数分计算时间匹配, 提高了速度. 计算可知, 优化后的算法共需要 12 次乘 32 次加.

3 2-D DCT IP 软核设计

总体上可以分为 4 个模块: 1-D DCT 模块, 矩阵转置模块, 控制逻辑, 和 MUX 选择器模块. 其结构如图 4 所示.

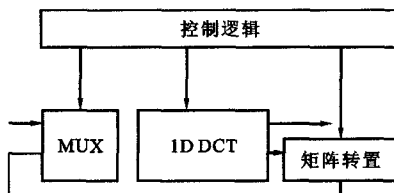


图 4 2-D DCT 结构图

3.1 一维 DCT 模块设计

由前面的分析可知, 二维离散余弦变换可以分解成两次一维余弦变换. 两次变换电路结构基本相同, 为节省面积, 采用一维 DCT 复用技术. 考虑到两次变换所处理的数据长度不同, 为增加通用性, 本模块输入位数为 11 位, 输出位数 12 位. 能够满足 MPEG1/2/4, JPEG 等视频标准对精度的要求.

由于算法中的 12 个乘法器系数固定且系数都是小数, 直接计算会引入浮点数, 处理麻烦而且消耗资源大. 本文在处理小数时, 在满足计算精度的前提下, 先将 12 个系数做放大处理, 转化为整数. 为减少芯片资源, 对所得整数采用更为精简的 CSD 编码^[5], CSD 编码是一种用最少的非零比特位来表示符号数的编码技术, 能够大幅减少“1”的数量, 提高乘法的效率.

例如: 系数 $f \times x$, f 二进制码表示为 1110100000, 当使用移位相加实现时用到 4 次移位,

3 次相加. 当 f 用 CSD 编码时, f 可以表示为 100(-1)0100000, 用到 3 次移位, 2 次相加. 可以看到采用 CSD 编码方式能大幅降低运算复杂度与延时. 表 2 中给出了 12 个系数 CSD 编码值.

表 2 12 个系数 CSD 编码值

系数	系数 $\times 1024$	CSD 码表示
C_6	196	11000100
$C_2 - C_6$	277	100010101
$-C_2 - C_6$	669	1010100(-1)01
a	108	1101100
b	743	1100(-1)0100(-1)
c	1112	10001011000
d	544	1000100000
e	326	101000110
f	928	100(-1)0100000
g	710	1011000110
h	141	10001101
i	426	110101010

在电路设计上, 充分利用 FPGA 的并行优势, 采用流水线结构并行处理图像数据. 8 点 1-D DCT 变换插入 8 级流水线, 即经过 8 个时钟完成一次变换. 提高了电路的速度.

为了进一步节省资源, 可以通过计算大体估算出电路输出的最大位数, 以确定每级流水线中寄存器的位数, 以减少面积的浪费.

3.2 矩阵转置的设计

矩阵转置相当于把第一次余弦变换的结果进行行列转换后作为二次余弦变换的输入. 相当于一个 88 的阵列. 常用的实现方法是使用普通的 RAM 实现, 完成一次转置需要 64 个时钟周期. 本设计采用状态机来实现矩阵的转置. 完成一次转置只需要 20 个周期大大提高了系统的运行速度. 其状态如图 5 所示, 分为 3 个态: 初始状态, 行存状态, 列取状态, 并用计数器来产生行列地址. 一个时钟周期可以完成 8 点数据的写操作, 一个时钟周期完成 8 点数据的读操作. 当执行写操作时, 每个有效信号来临时计数器执行加 1 操作, 产生行存地址. 当计数器加到 7 时, 表示 64 个数据存储完毕, 进入列取状态, 此时计数器开始执行减 1 操作, 用于产生列取地址, 直到计数器减到 0, 表示一次转置完成, 等待下一次数据的来临. 电路实现上用 8 片 11bit8 RAM 组成变换矩阵. 采用这种方法实现了一维离散余弦变换模块与矩阵转置模块在传输速率上的匹配. 整个矩阵转置模块共需 820 个逻辑单元.

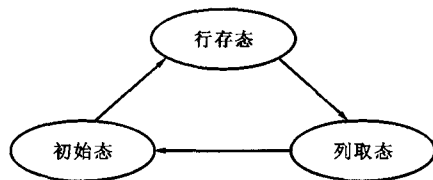


图5 矩阵转置状态图

3.3 控制模块的设计

为了使各个模块有序的工作,控制电路用有限状态机来实现。由于数据是串行输入的,在到达 DCT 变换模块之前必须先经过串并转换模块进行并行化处理。控制模块产生的控制信号用于控制串并转换模块, DCT 变换模块的使能端,和 MUX 选择器模块控制端。当数据来临后, DCT 变换模块使能,串并转换模块使能,通过 MUX 选择器将串并转换模块与 DCT 变换模块相连,进行行的 DCT 变换。经过 64 个时钟完成了 64 个数据的输入后,选择器将串并转换模块与 DCT 变换模块断开,将转置矩阵模块的输出作为 DCT 模块的输入,进行列的 DCT 变换。28 个时钟后,完成了一次 2-D DCT 的变换,等待下次处理数据的来临。

采用自顶向下的设计方法,用 Verilog HDL 硬件描述语言对上述模块进行了行为级描述,设计出了基于 Loeffler 算法的 2-D DCT IP 软核。

4 测试与验证

整个 IP 软核,在 QuartusII 8.0 设计工具中完成编译、综合、布局布线和后仿真。所采用的芯片为 Altera 公司的 EP2C35F672C8,图 6 为综合后显示的系统所占用的硬件资源情况。可以看到整个 IP 软核共需 2565 个逻辑单元,占用资源较少。

Revision Name	dct
Top-level Entity Name	DCT2
Family	Cyclone II
Device	EP2C35F672C8
Timing Models	Final
Total logic elements	2 565/33 216 (8%)
Total combinational functions	1 741/33 216 (5%)
Dedicated logic registers	2 398/33 216 (7%)
Total registers	2 398
Total pins	107/475 (23%)
Total virtual pins	0
Total memory bits	36/483 840 (<1%)
Embedded Mul tiplier 9-bit elements	0/70 (0%)
Total PLLs	0/4 (0%)

图6 综合结果

FPGA 模拟显示 IP 软核最高工作频率可以达到 139.43 MHz。比文献[6]中报道的提高了近 1 倍,大大缩小了图像压缩时间。如果想要获得更高的工作频率可以从两方面改进:(1)电路采用专用集

成电路(ASIC)来实现;(2)适当增加电路流水线的级数,当然增加流水线的级数必然会带来占用芯片面积的增加。这就需要速度和面积均衡考虑。

文中采用将输出结果矩阵与在 Matlab 中的仿真结果做比较,来验证算法的正确性。测试输入采用《Image and Video Compression Standard》第二版给出的测试矩阵作为 DCT 模块的输入数据。矩阵 RAW 即为输入测试矩阵。

$$RAW = \begin{bmatrix} 40 & 33 & 33 & 22 & 40 & 40 & 36 & 26 \\ 43 & 26 & 33 & 22 & 43 & 43 & 22 & 36 \\ 43 & 40 & 19 & 36 & 33 & 33 & 15 & 26 \\ 36 & 43 & 26 & 33 & 29 & 29 & 19 & 4 \\ 33 & 33 & 29 & 26 & 15 & 33 & 26 & 4 \\ 36 & 33 & 33 & 26 & 22 & 29 & 26 & 12 \\ 33 & 40 & 29 & 26 & 33 & 12 & 4 & 26 \\ 33 & 29 & 22 & 12 & 12 & 4 & 8 & 0 \end{bmatrix}$$

矩阵 B 是在 QuartusII 8.0 中输出的结果矩阵。

矩阵 C 是 Matlab 取整后的运算结果矩阵。

$$B = \begin{bmatrix} 215 & 45 & -3 & 21 & -4 & -9 & 2 & -2 \\ 39 & -22 & -3 & 15 & 3 & -3 & 9 & 1 \\ -5 & -8 & 8 & -6 & 8 & -10 & 5 & 15 \\ 13 & -10 & 0 & 1 & -13 & 8 & 5 & 9 \\ -16 & 7 & -3 & 4 & -18 & 12 & -9 & 2 \\ 8 & 6 & -3 & -6 & 7 & -10 & -9 & -2 \\ -3 & 2 & 3 & 4 & -6 & 10 & -5 & -3 \\ -1 & -2 & 2 & -3 & 4 & -3 & -1 & -6 \end{bmatrix}$$

$$C = \begin{bmatrix} 217 & 45 & -1 & 21 & -3 & -9 & 2 & -2 \\ 39 & -22 & -3 & 15 & 3 & -3 & 8 & 0 \\ -4 & -7 & 9 & -4 & 9 & -10 & 5 & 15 \\ 13 & -11 & -1 & 0 & -13 & 7 & 5 & 8 \\ -16 & 7 & -2 & 5 & -17 & 12 & -8 & 2 \\ 8 & 6 & -4 & -6 & 7 & -10 & -10 & -2 \\ -3 & 2 & 3 & 4 & -5 & 10 & -5 & -4 \\ -1 & -1 & 3 & -3 & 4 & -3 & -1 & -6 \end{bmatrix}$$

所设计的 IP 软核仿真结果与 Matlab 中运算结果很接近,验证了算法的正确性。可以看出两者并不完全相同,说明 IP 软核在进行 DCT 变换时存在一定的误差,这是因为在右移操作时产生的截断误差和系数量化时产生的量化误差造成的。

5 结束语

文中以设计实现适用于 JPEG 及 MPEG 视频压缩协议的高速编码电路为目的,设计出了基于 Loeffler 算法的 2-D DCT IP 软核。通过用移位和加法运算代替乘法运算,并对乘法系数采用 CSD 编

(下转第 144 页)

分析的可行性与有效性.

参考文献:

- [1] Paul Kocher. Timing attacks on implementations of diffie-hellman, RSA, DSS, and other systems[C]// Proceedings Crypto '96, LNCS. New York: Springer-Verlag, 1996:104-113.
- [2] 邓高明, 赵强, 张鹏, 等. 针对密码芯片的电磁频域模板分析攻击[J]. 计算机学报, 2009, 32(4): 602-610.
- [3] 邓高明, 张鹏, 赵强, 等. 针对密码芯片的电磁模板分析攻击[J]. 微电子学与计算机, 2010, 27(1): 1-4.
- [4] Martin Goldack. Side-channel based reverse engineering for microcontrollers [D]. Bochum: Ruhr-University, 2008.
- [5] Dennis Vermoen. Reverse engineering of Java card applets using power analysis [D]. Netherlands: TU Delft, 2007.
- [6] Jan M Rabaey, Anantha Chandrakasan, Borivoje Nikoli. 数字集成电路—电路、系统与设计[M]. 2 版. 周润德,

译. 北京: 电子工业出版社, 2004: 141-166.

- [7] SungMo Kang, Yusuf Leblebici. CMOS 数字集成电路—分析与设计[M]. 3 版. 王志功, 窦建华, 译. 北京: 电子工业出版社, 2005: 97-98.
- [8] 褚杰, 赵强, 丁国良, 等. 密码系统差分功耗分析攻击及功耗模型[J]. 华中科技大学学报: 自然科学版, 2008, 36(6): 18-21.

作者简介:

- 邓高明** 男, (1983-), 博士研究生. 研究方向为电磁信息检测与主动防护技术.
- 吴恒旭** 男, (1982-), 硕士研究生. 研究方向为信息安全.
- 张鹏** 男, (1976-), 博士研究生. 研究方向为电磁信息检测与主动防护技术.
- 邹程** 男, (1982-), 博士研究生. 研究方向为电磁信息检测与主动防护技术.
- 陈开颜** 女, (1970-), 博士, 硕士生导师. 研究方向为信息安全.

(上接第 139 页)

码, 1-D DCT 复用等措施来减少芯片面积, 通过优化电路结构, 加入 8 级流水线, 优化转置矩阵来提高电路速度. 仿真结果验证了算法的正确性. 最高工作频率可以达到 139.43MHz. 能够满足视频图像压缩的实时性要求.

参考文献:

- [1] Ahmed N, Natrajan T, Rao K R. Discrete cosine transform[J]. IEEE Transaction on Computers, 1974(C-23): 90-95.
- [2] Agostini L V, Saraiva I. Pipelined fast 2-D DCT architecture for JPEG image compression[C]// SBCCI'01, the 14th Symposium on Integrated Circuits and Systems Desing. Brazil: Pirenopolis, 2001: 226-231.
- [3] 王继东, 张会生, 范忠亮. 基于 NEDA 算法的二维 DCT

硬件加速器的设计实现[J]. 微电子学与计算机, 2008, 25(5): 166-168.

- [4] Loeffler C, Lightenberg A, Moschytz G. Practical fast 1-D DCT algorithms with 11 multiplications [C]// ICASSP'89, Proceedings of IEEE International Conference on Acoustics Speech and Signal Processing. USA: New York, 1989, 2(5): 988-991.
- [5] Uwe Meyer Baese. 数字信号处理的 FPGA 实现[M]. 刘凌, 译. 北京: 清华大学出版社, 2006: 32.
- [6] 陈伟, 卢贵主, 郑灵翔. 二维 DCT 算法及其精简的 VLSI 设计[J]. 微电子学与计算机, 2008, 25(3): 108-110.

作者简介:

- 郭宝增** 男, (1953-), 教授, 硕士生导师. 研究方向为集成电路设计.