

文章编号: 1000-8829 (2004) S0-0198-03

# H.264 变换编码的 DSP 实现

## An Implementation of Transform Encoding on DSP in H.264

(西北工业大学 电子信息学院, 陕西 西安 710072) 沈海涛, 樊养余, 王凤琴, 郝重阳

**摘要:** H.264 是新一代的视频压缩标准。本文首先研究了该标准并着重讨论了其关键技术整数变换编码算法, 然后简要介绍了 TMS320C6416 DSP 的出色性能及其软件流水线原理, 最后结合 6416 DSP 的指令特点提出了一种编程方法, 并写出了该变换编码算法的高效汇编代码, 为实时视频信号压缩提供了保证。

**关键词:** H.264; 整形变换; DSP

**中图分类号:** TN919. 8

**文献标识码:** A

**Abstract:** H.264 is the newest video compression standard. This standard is discussed in this paper, especially its key technique-integer transform code arithmetic. The excellent performance and the principle of pipeline of TMS320C6416 DSP is then introduced simply. We put forward a new program method combine with the special instructions of 6416 DSP, and write out the efficiently ASM program code, which provides guarantees for real-time process of video signal.

**Key words:** H.264; integer transform; DSP

由于具有快速算法, 易于硬件实现, 且其编码性能良好的特点, DCT 变换在 MPEG-x 和 H.26x 两大视频编码标准中得到了广泛的应用。在 JVT (Joint Video Team) 制定的最新的编码标准 H.264 (或称 AVC 或 MPEG-4 part10) 中, 提出了  $4 \times 4$  块的整形变换, 对残差进行基于块的变换编码, 由于变换是整数操作而不是实数运算, 因此避免了“反变换误差”, 其过程和 DCT 基本相似。

本文主要介绍 H.264 整形变换的基本原理及其在 TI TMS320C6416 DSP 芯片上的实现。TMS320C6416 定点 DSP 是一台主频率为 600 兆赫 (MHz) 的器件, 其性能高达每秒 48 亿次指令。处理密集型 C6416 DSP 架构采用了二级高速缓存以及每周期能够并行执行 8 个 32 位指令的超长指令字 (VLIW), 适合于实时图像信号处理。

### 1 H.264 标准简介

H.264 标准压缩系统由视频编码层 (Video Coding Layer, VCL) 和网络提取层 (Network Abstraction Layer, NAL) 两部分组成。VCL 中包括 VCL 编码器与 VCL 解码器, 主要功能是视频数据压缩编码和解码, 它包括运动补偿、变换编码、熵编码等压缩单元。NAL 则用于为 VCL 提供一个与网络无关的统一接口, 它负责对视频数据进行封装打包后使其在网络中传送。它采用统一的数据格式, 包括单个字节的包头信息、多个

字节的视频数据与组帧、逻辑信道信令、定时信息、序列结束信号等。包头中包含存储标志和类型标志。存储标志用于指示当前数据不属于被参考的帧。类型标志用于指示图像数据的类型。VCL 可以传输按当前的网络情况调整的编码参数。

H.264 标准的编码框架与先前的视频编解码标准一致, 即基于块的混合编码方式, 如图 1 所示。

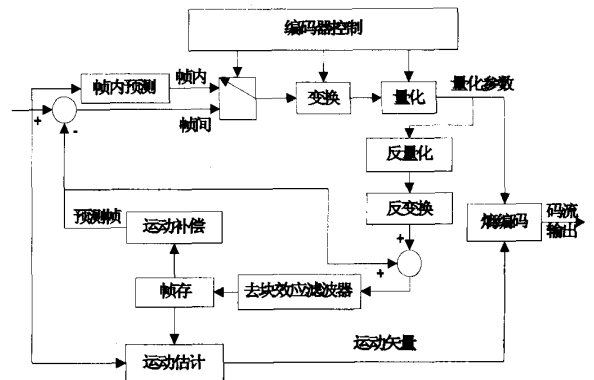


图 1 H.264 结构框图

与先前的标准相似, H.264 对残差采用基于块的变换编码, 而且这个变换是整数操作, 而不是实数运算, 其过程和 DCT 基本相似。这种方法的优点在于: 在编码器和解码器中允许精度相同的变换和反变换, 便于使用简单的定点运算方式。也就是说, 这里不存在浮点运算引起的“反变换误差”。变换的单位是  $4 \times 4$  块, 而不是以往常用的  $8 \times 8$  块。由于用于变换块的尺寸缩小, 运动物体的划分更精确, 这样, 不但变换计算量比较小, 而且在运动物体边缘处的衔接误差也大为减小。

为了使小尺寸块的变换方式对图像中较大面积的平滑区域不产生块之间的灰度差异, 可对帧内宏块亮度数据的 16 个  $4 \times 4$  块的 DC 系数 (每个小块一个, 共 16 个) 进行第二次  $4 \times 4$  块的变换, 对色度数据的 4 个  $4 \times 4$  块的 DC 系数 (每个小块一个, 共 4 个) 进行  $2 \times 2$  块的变换, 如图 2 所示, 其中:

$4 \times 4$  变换 1:  $4 \times 4$  DCT 变换的整形近似;

$4 \times 4$  变换 2: 所有元素均为  $\pm 1$  的 Hadamard 变换。

整形变换核定义如下:

$$H = \text{round}\left(\frac{5}{2} H_{\text{DCT}}\right) = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & -2 \\ 1 & -1 & -1 & 1 \\ 1 & -2 & 2 & -1 \end{bmatrix}$$

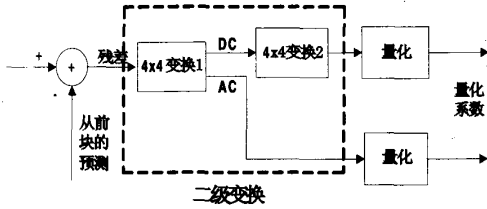


图 2 H.264 变换与量化

由变换公式  $Y=H \times H^T$  即可推得  $4 \times 4$  整形变换的蝶形结构:

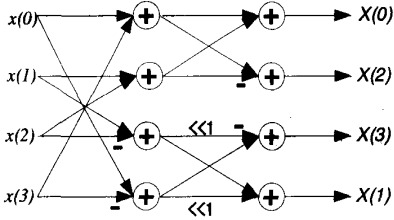


图 3 正变换

由于该整形变换算法避免了  $4 \times 4$  DCT 中由浮点运算带来的数据舍入误差, 因此它不存在反变换误差, 而且更容易在定点 DSP 上实现。

## 2 DSP 性能简介

TMS320C6416 为定点 DSP 芯片, 它具有两个数据通道, 每个数据通道包含 4 个功能单元和 32 个通用寄存器, 4 个功能单元分别为 S、L、M 和 D, 寄存器组 A 可以通过交叉通道与另一个寄存器组 B 的功能单元相连, 交叉通路允许一侧数据通道的功能单元访问另一侧寄存器组的 32 位操作数, 控制寄存器组用来设定和控制不同处理器的操作。

并行处理是其突破传统设计而获得高性能的关键, 6416DSP 可以在一个指令周期并行执行 8 条 32 位指令。所以在 600 MHz 的时钟频率下, 其处理能力可以达到 4 800MIPs, 如此的计算速度, 完全可以实现数字信号的实时处理。

任何指令的处理都可以分成几个子操作, 每个操作由不同的功能单元来实现, 对每个功能单元来说, 每隔一个时钟周期就可以进入一条新的指令, 这样在同一时间内, 在不同的单元中可处理多条指令, 这种方式称为“流水线”(pipeline)工作方式。软件流水线是用来安排循环指令的, 是循环的多次迭代并行执行的一种技术。图 4 所示为一个软件流水线循环示意图。图中各次迭代由 A、B、C、D 和 E 代表, 其后的数字代表各次迭代的是第几条指令。同一个周期内最多可以执行 5 次迭代的的不同指令, 即图 4 中阴影部分所示, 这一部分也就是循环的核心部分, 在核心代码中, 5 次迭代的的不同指令并行执行。核心前面的部分称为循环填充 (prolog), 核心后面的部分称为循环排空 (epilog)。

6000 系列 DSP 的指令具有延迟间隙, 即从指令的源操作数读取到执行的结果可以被访问所使用的指令周期, 如数据读取指令 (LDW) 在第  $i$  周期执行, 在第  $i+4$  个周期时数据写入目标寄存器, 从第  $i+5$  个周期开始才可以对这个数据进行处理。

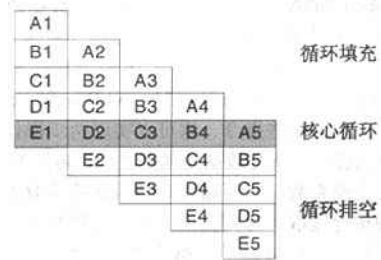


图 4 软件流水线

## 3 整形变换的 DSP 汇编实现

只要适当的安排代码的循环填充和循环排空部分以及循环的迭代间隔, 就可以编写出尽可能高效的汇编代码。以下是  $4 \times 4$  整形变换的部分核心代码 (数据在存储器中按块按行存放):

```

loop1:
    LDDW .D1T2 *-A_i[1],B_f0s:B_f0f
    || LDDW .D2T1 *+B_i[1],A_f3s:A_f3f
    .....
    LDDW .D1T1 *A_i++[4],A_f1s:A_f1f
    || LDDW .D2T2 *B_i++[4],B_f2s:B_f2f
    || BDEC .S2 loop1, B1
    ||[B0]MPY .M2 B0, 4, B0
    .....
    PACK2 .L1 A_F1f, A_F0f, A_F0nf
    || PACKH2 .S1 A_F1f, A_F0f, A_F1nf
    || PACKH2 .S2 B_F3s, B_F2s, B_F3ns
    || PACK2 .L2 B_F3s, B_F2s, B_F2ns
    .....
    ADD2 .L2X B_f0f, A_f3f, B_p0
    || ADD2 .L1X A_f1f, B_f2f, A_p1
    || SUB2 .S2X B_f0f, A_f3f, B_r0
    || SUB2 .S1X A_f1f, B_f2f, A_r1
    .....
    [!B0]STDW .D1T1 A_F1ns:A_F1nf,*-A_i[8]
    ||[!B0]STDW .D2T2 B_F3ns:B_F3nf,*-B_i[7]
    .....
    
```

对  $4 \times 4$  的数据块进行二维整形变换, 可以采用与 DCT 变换相类似的方式, 即先列后行。每次从存储器中读取一列数据, 分别存放到不同的通用寄存器中, 处理完之后再将结果返回到存储器中, 然后开始下一列数据的处理。列变换完后对行进行变换, 从而完成对全部数据块的二维变换。

由于数据在存储器中是按 32 位存放的, 而实际上每个像素点值的范围不会超过 10 位, 而且经过一些简单的加减、移位操作后也不会超出 16 位的范围, 所以可以对所有的数据进行打包压缩, 这样既节省了存储空间, 又可以充分利用 6416DSP 的高效指令, 从而提高程序的运行速度。

由于 LDDW 可以一次读取 64 位数据, 因此对数据进行打包后, 该指令就可以读取存储器中的 4 个数, 也就是  $4 \times 4$  块的

一行,用4条LDDW指令就可以读取整个数据块,然后用ADD2, SUB2等指令对数据进行处理,如图5所示。其中A17与A18的高16位分别是x01、x31,低16位分别是x00、x30,执行ADD2指令后, A19H=x01+x31; A19L=x00+x30, SUB2指令与此类似。因为这些指令一次可以实现两步操作,因此处理速度成倍提高。所有数据处理完之后,再用STDW指令存储处理完的变换系数,其存储方式与数据的读取方式一致。很显然,这种按行读取、对列变换的方法使得数据的处理速度大大提高。对于一个4×4块的数据块的列变换,仅需要8个时钟周期就可以完成。

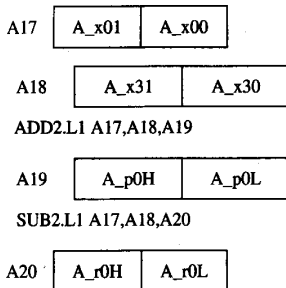


图5 6416DSP汇编指令

对数据作行变换时,由于DSP系统资源的限制以及数据结构的变化而不能达到与列变换同样高的效率。为了解决这个问题,笔者提出了一种行读取、列变换、行存储的方法。其原理是在列变换后,将列变换系数按行存放在原数据块处,即对原列变换系数矩阵作转置。而后的行变换就转变成对列变换系数矩阵的列变换,其程序就与列变换完全一样了。这种方法巧妙的将行变换转化为列变换,充分的利用了DSP资源,降低了代码复杂度,提高了数据的处理速度。

程序中设置了A<sub>i</sub>和B<sub>i</sub>两个读/写指针以加快数据的读写速度;用BDEC实现跳转,节省了循环次数自减指令,使程序更加简洁;利用PACK2和PACKH2等指令对数据进行重组,实现对数据矩阵的转置。

经过恰当的安排流水线编排,使L、S和D功能单元得到了最大限度的利用。测试发现,用汇编语言实现的整形变换比用C语言实现的变换效率提高很多。用CCS对C语言程序进行多次软件优化后,每个4×4块的正变换需要18(列变换)+24(行变换)共42个CPU周期,而汇编程序仅需要8(列变换)+8(行变换)共16个CPU周期,代码执行效率提高近62%。

完成对残差的4×4整形变换后,修改程序的一小部分,就可以实现对DC系数的Hadamard变换,从而完成全部的变换过程,准备开始量化。

## 4 结论

H.264优越性能的获得不是没有代价的,其代价是计算复杂度的大大增加;变换/反变换时,视频图像压缩系统中的关键模块运行速度及精度对整个系统的处理性能影响很大。因此将该模块移植到高速的定点DSP芯片上实现,在对提高系统性能上的重要意义是不言而喻的。目前,大部分数字通讯系统都是基于DSP的,如果能用DSP芯片完成视频图像处理将会有助

于数字通讯系统的集成。我们在DSP芯片进行的这一有益的尝试,为后续的工作打下了良好的基础。

### 参考文献:

- [1] Iain E G Richardson. H.264/MPEG-4 Part 10 White Paper [EB/OL]. www.vcode.com, 2003-03.
- [2] 周霖. DSP信号处理技术应用[M].北京:国防工业出版社, 2004-01.
- [3] Texas Instruments. TMS320C6000 CPU and Instruction Set Reference Guide[M]. 2000-08.
- [4] Henrique Malvar. Low-Complexity Transform and Quantization in H.264[M]. 2003-02.
- [5] 视频编码新标准H.264的技术亮点[EB/OL]. www.videosky.com.
- [6] Ali Jerbi. H.264 MPEG-4 AVC Video Coding Standard[EB/OL]. www.ubvideo.com, 2003-11.

作者简介:沈海涛,男,西北工业大学电子信息学院推荐硕士研究生,专业为电子信息工程。

(上接第197页)

其中快截距 $g_0$ 有8种可选取值,其他3个参数有4种可选取值,组合起来共有512种声音掩蔽模型。比特分配监控模块的一个任务就是选择一组最佳的模型,使得计算出来的核心掩蔽曲线最逼近理想比特分配曲线。

比特分配监控模块的另一任务就是调整核心掩蔽门限的偏移量来获得噪声曲线,再根据信号和噪声曲线计算信噪比。比特分配是根据信噪比的大小进行分配的。比特分配监控模块对信噪比的计算与总的可用于尾数编码的比特数密切相关。比特分配监控模块根据目前可用于尾数分配的总比特数对核心掩蔽曲线进行上下平移,所得到的结果作为噪声门限。根据噪声门限和信号的功率谱密度进行比特分配。当目前使用的比特数大于比特池中的可用比特数时需要调大噪声门限。反之,则可以将噪声门限调小。比特分配监控模块通过调整核心掩蔽曲线的偏移量也就是调整噪声门限的方法,最终使比特池中的可用比特数得到最大的利用。

## 2 总结

AC-3音频标准中,比特分配模块在很大程度上决定了编/解码器的质量。本文对AC-3标准中比特分配的核心算法进行了深入研究,为AC-3音频编/解码器的设计与实现奠定理论基础。

### 参考文献:

- [1] TSC. Digital Audio Compression(AC-3)[S]. ATSC Standard. 1994.
- [2] C. Todd, G. Davidson, M. Davis, etc.. AC-3: Flexible Perceptual Coding for Audio Transmission and Storage[A]. AES 96th Convention, Preprint. Amsterdam, 1994.

作者简介:张涛,男,天津大学电子信息工程学院信号与信息处理专业博士生。主要研究方向为DSP应用, HDTV。