文章编号:1002-8692(2008)12-0025-04

基于DSP平台的AVS视频编码器设计优化 ·实用设计·

张 妍, 薛永林, 赵 康
(清华大学 数字电视技术研究中心, 北京 100084)

【摘 要】介绍了在TMS320DM6446 DSP平台上实现AVS视频编码器的算法设计与优化方法。在软件整体设计优化的基础上,重点对运动估计等算法进行了优化改进;同时针对平台特点给出结构优化方法,主要包括提高代码并行性及存储器和数据搬移的优化。测试结果表明,通过优化,在保证图像质量损失较小的情况下,编码器的编码速率有显著提高。

【关键词】AVS标准;TMS320DM6446;算法优化;结构优化
【中图分类号】TN919.8 【文献标识码】A

Algorithms Design and Optimization of AVS Video Encoder Based on DSP

ZHANG Yan, XUE Yong-lin, ZHAO Kang
(Digital Television Technology R&D Center, Tsinghua University, Beijing 100084, China)

【Abstract】 In this paper, algorithms design and optimization of a AVS video encoder using TMS320DM6446 DSP of Texas Instruments are presented. On the basis of the overall design and optimization of software, algorithms such as motion estimation are mainly improved. According to the architecture of TMS320DM6446, the architecture level optimizations are used, which involves software pipeline and memory access efficiency. The test results show that the coding rate is dramatically improved, while the performance of video sequence is considerably good.

【Key words】 AVS; TMS320DM6446; algorithm optimization; architecture optimization

1 引言

AVS[1]是我国具有自主知识产权的视频压缩编码标准,采用DSP实现AVS视频实时编码具有开发周期短、灵活性强等优点,是适合当前产业发展的一个重要途径。本文所述系统采用TI公司最新推出的数字媒体处理器TMS320DM6446为开发平台,给出了AVS视频编码器的软件设计方案和算法优化方法,并针对平台特点给出结构优化方法。

2 AVS视频编码器

与传统的视频编码标准相同,AVS也采用基于块的混合编码框架。主要模块包括帧内模式选择、帧内预测、运动估计、运动补偿、DCT变换、量化、反量化、IDCT变换、VLC编码等模块。在算法设计上,为了便于运动搜索,1/2像素插值采用整帧插值,即编码完每一帧后,对非B帧进行1/2像素插值并储存插值结果,而1/4像素插值在运动估计时即时进行,因此1/2像素插值可以作为一个单独的模块考虑。

编码器的计算量主要集中在运动估计和运动补偿,1/2像素插值也占用了较多的时间,DCT模块和量化模块实际上在本系统中并不是单独作为一个模块存在的,而是在帧内模式选择、帧内预测和运动补偿模块中都有调用,因此减少DCT和量化的运算量对提高系统编码速率也有重要作用。

3 编码器优化

视频编码器优化可以分为算法优化和结构优化。算法优化与平台无关,仅对C代码进行优化,使之减少运算复杂度;结构优化则要根据开发平台的特点进行优化,使编码器充分利用硬件资源,主要包括提高处理器的访问效率和代码的并行性。

3.1 算法优化

运动估计是整个系统中运算量最大的模块,是算法优化的重点。图1所示为运动估计算法流程,其中的最佳匹配采用绝对误差和作为判断准则,即SAD (Sum of Absolute Difference)最小的块为当前宏块的最佳匹配。运动估计优化的基本思想是尽量减少对系统性能影响较小的计算。本系统对图1中所示的带阴影模块进行了以下几个方面的优化。

1) 改进可用预测模式选择

AVS的帧间编码宏块有4种可选预测模式:16×16,16×8,8×16和8×8。对测试序列进行统计发现,最佳模式

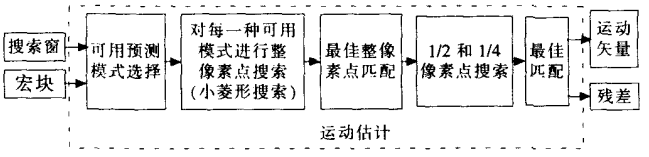

图 1 运动估计流程图

为 16×8、8×16 两种过渡模式的宏块很少，且具有空间相关性，因此可以在运动搜索之前判断这两种模式是否可用，减少不必要的计算。

根据空间相关性，把预测模式分为 3 种情况，即图 2 所示的两种情况和其他情况。对于情况一，如果当前宏块 C 上方的宏块 T 最优模式为 8×16，即物体边界垂直穿过上方宏块，则当前宏块也很有可能包含垂直边界；对于情况二，如果左侧宏块 L 的最优模式为 16×8，则当前宏块也很有可能包含水平边界。3 种情况的可用模式如表 1 所示，这种算法可以将帧间模式选择的计算复杂度降低约 30%，而 *PSNR* 仅降低约 0.04 dB。

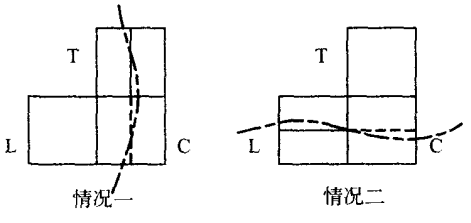

图 2 相邻宏块模式的相关性示意图

表 1 不同情况的可用预测模式

情况	可用预测模式
一	16×16,8×16,8×8
二	16×16,16×8,8×8
其他	16×16,8×8

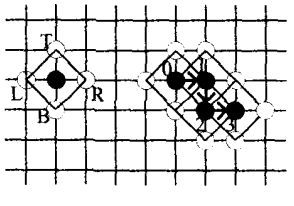

图 3 小菱形搜索

2）改进小菱形搜索

图 3 所示为小菱形搜索的示意图，对当前整像素点的 T,B,L,R 等 4 个像素点分别计算 *SAD*，*SAD* 最小的点作为新的搜索中心继续搜索，直到 T,B,L,R 等 4 个点的 *SAD* 都大于中心点 *SAD* 为止。除第一步小菱形搜索的每一步搜索由于出现了搜索点的覆盖，最多只需要计算相邻 3 个点的 *SAD*。为减少菱形搜索的计算量，同时不增加额外的存储，本系统采用一步路径记录法，用变量 *path* 记录路径，具体步骤如下：

（1）搜索 T,B,L,R 等 4 个像素点，并记录搜索结果，如 R 对应的 *SAD* 最小，则 *path=R*。

（2）根据 *path* 判断当前需要搜索的点，如果 *path=R*，则当前只需搜索 T,B,R（当前点的 L 即第一步搜索的 R），再记录搜索结果。

（3）重复（2），直至搜索结束。

改进的小菱形搜索减少约 1/4 的菱形搜索计算量，同时不会影响 *PSNR*。

3）改进 1/2 和 1/4 像素点搜索[2]

通过小菱形搜索找到运动估计的最佳整像素点后，需要进行 1/2 和 1/4 像素点搜索。通常的做法是对整像素点周围 8 个 1/2 像素点计算 *SAD*，找到最佳 1/2 像素点，再对 1/2 像素点周围 8 个 1/4 像素点进行搜索。如图 4所示，每一次搜索都要计算 L,R,T,B,LT,RT,LB,RB 等 8 个点的 *SAD*[3]。为减少计算量，考虑到 L,R,T,B 等 4 个像素点相比 4 个角像素点距中间像素的空间要小些，本系统采用两步的改进搜索算法：首先计算 L,R,T,B 等 4 个像素点的 *SAD*，找到最佳像素点；如果最佳像素点是中心像素点，则直接进行下一步搜索，否则计算最佳像素点相邻 2 个角的 *SAD*。例如第一步搜索找到 R 是最佳像素点，第二步只搜索 RT 和 RB，而 LT 和 LB 则不再进行搜索。这种改进的搜索算法可以减少 25%的 *SAD* 计算量，对于即时插值的 1/4 像素点搜索还可以减少 25%的 1/4 插值计算量。用不同的测试序列进行测试，采用改进的 1/2 和 1/4 像素点搜索算法，*PSNR* 变化均未超过±0.02 dB。

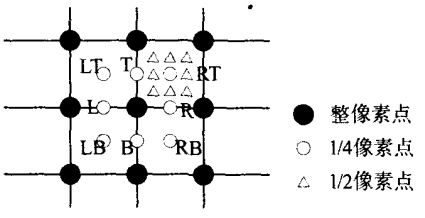

图 4 1/2 和 1/4 像素点搜索

4）域值法

域值法的核心思想是搜索到足够好的匹配块后就停止搜索，这种算法是以牺牲一定的 *PSNR* 为代价，获得更快的编码速度。在进行运动估计前设置好 *SAD* 域值 SAD_{th}，算法进行小菱形搜索、1/2 像素搜索或者 1/4 像素搜索时，每一步计算得到的 *SAD* 只要小于 SAD_{th} 就停止搜索，将当前模式设为最佳模式，当前运动矢量设为最佳运动矢量。域值法的关键是 SAD_{th} 的选取，SAD_{th} 过大，*PSNR* 将严重降低，过小起不到较好的省时效果。这里选取量化步长的整数倍作为搜索的域值，即 $SAD_{th}=T_1\times Q_P$[2]。用不同的 T_1 对 Foreman 进行测试，可得表 2。可见随着 T_1 的变大，*PSNR* 稍有降低，编码速度升高，可根据需要选取 T_1 值。

3.2 结构优化

TMS320DM6446 为 ARM9+C64x 的双核处理器结

表 2 T_1 对 *PSNR* 和编码速度的影响(Q_P=30)

T_1	*PSNR*/dB	编码速度/(f/s)	码率/(Kbit/s)
0	37.075	8.190	899.326
20	37.049	9.740	900.231
40	36.862	13.520	1 003.161
60	36.762	16.200	1 114.757

构[3]。其核心处理器 TMS320C64x DSP 主频高达 594 MHz，具有丰富的专门为多媒体运算优化的指令集。TMS320C64x DSP 主要特点是采用了超长指令字 VLIW (Very Long Instruction Word)结构[4]，在 VLIW 处理机中，最多可有 8 个功能单元并行工作。提高 VLIW 核内的代码并行性和 DSP 片上存储器的访问效率是结构优化的重点。

3.2.1 提高代码并行性

提高代码并行性主要从以下两个方面进行：

1) C 代码的优化

TMS320C64x DSP 支持标准 C 语言，并对标准 C 语言进行了补充和扩展，增加了一些供 C 语言使用的内联函数(intrinsics)，内联函数是直接映射为内联的编译器指令的特殊函数，如 _mpyhl，_add2 等，C 语言可以直接调用[5]。当对一连串短型数据进行操作时，可使用内联函数对这些数据进行操作，从而减少对内存的访问，如使用 _add2 内联函数对 2 个连续短型数据进行相加运算。另外在 VLC 模块的哥伦布指数编码中，需计算每个码字的编码级数，传统的做法是从零级逐个查找，但效率较低，而采用 _lmbd 内联函数则可以很快地判断出级数不需循环查找；_lmbd 还可以用来计算码字长度；编码数据写入码流时，采用 _sshvl 内联函数进行数据的移位操作也可提高编码效率，减少不必要的函数调用[5]。

在 C 代码中，对指针变量增加 restrict 和 const 修饰符可以帮助编译器确定哪些指令不相关，安排它们并行执行。同时利用编译优化选项可以帮助编译器创建合适的软件流水。

2) 线性汇编优化

C 代码优化一般不能满足系统性能的需求，需要抽出 C 代码中的低效率段用线性汇编重新编写。线性汇编是一种简化的汇编语言，线性汇编不需给出寄存器、并行与否、延迟周期等，汇编优化器会根据代码情况确定这些信息。线性汇编比较典型的思路是使用字访问短型数据或使用双字访问字和循环展开。对大量的块数据进行处理的模块，如 DCT、量化、像素插值等，使用取字指令(LDW)或取双字指令(LDNDW)同时读取多个数据，可以大大减少数据访问时间。循环展开目的是在一个循环中同时处理多个数据，减少循环内的周期数，例如使用 SUBABS4 和 DOTPU4(*SAD* 计算)，AVGU4(插值)等指令同时处理 8 个像素值。本系统分别对 DCT、量化、帧内预测、1/2 像素插值、1/4 像素插值、计算 *SAD*、计算残差、重建、VLC 编码数据处理等模块进行了线性汇编优化，优化后主要函数的占用指令周期数对比如表 3 所示，可见通过线性汇编优化，所有模块的效率均有提高，部分模块的速度可以提高 10~20 倍。

表 3 线性汇编优化前后主要函数指令周期对比

函数名称	C指令周期/个	线性汇编指令周期/个
DCT	16 755	8 168
IDCT	13 995	7 866
量化	17 931	2 346
反量化	1 225	1 054
1/2 像素水平插值	248 052	45 719
1/2 像素垂直插值	765 251	31 301
1/4 像素水平插值(16×16)	36 549	8 970
1/4 像素垂直插值(16×16)	87 525	6 908
SAD(16×16)	14 103	1 428
残差(亮度块)	19 250	2 059
重建(亮度块)	22 337	2 130

3.2.2 存储器和数据搬移优化

TMS320C64x DSP 内核具有二级片内存储器结构[4]：第一级由 32 Kbyte 程序缓存(L1P)和 80 Kbyte 数据缓存(L1D)组成；第二级存储器(L2)容量为 64 Kbyte，用于存储程序和数据。如图 5 所示，每块片内存储空间都可以配置 RAM 和 Cache 的大小，CPU 读取 L1，L2 和片外内存的速度分别为 600 Mbit/s、300 Mbit/s 和 100 Mbit/s，充分利用片内存储器可以有效地提高编码速度。

本系统将片上缓存尽可能大的配置为 Cache，以提高 CPU 的处理速度。L1D 中的 48 Kbyte 空间配置为 RAM，用于存储经常访问的全局数据，以及 stack，data，const 等字段，占用时间比较大的模块处理的数据也放在 RAM 上，如 DCT 变换处理的 64 个短型数据等。

TMS320DM6446 还具有 64 个独立的 EDMA 通道，可在 CPU 后台进行大量的数据搬移，提高访问效率。在本系统中对于 B 帧图像，需要先缓存，处理完后面的 I 帧或 P 帧再处理 B 帧，缓存 B 帧的过程可以通过 DMA 实现，节省 CPU 操作的时间。

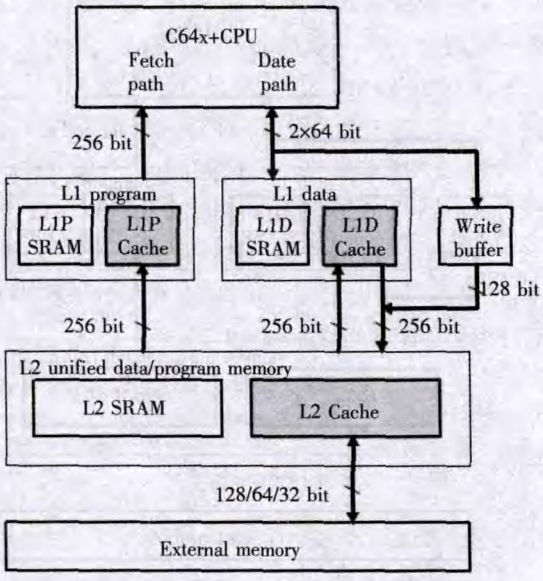

图 5 TMS320C64x DSP 缓存结构

4 编码器性能测试

本编码系统在 TMS320DM6446 评估板上对标准 CIF 序列(352×288)进行编码测试,结果如表 4 所示。可见本系统所做的算法优化在平均 *PSNR* 损失较小(小于 0.11 dB)、码率增加较小(小于 5%)的情况下,可将编码速度提高一倍;结构优化对编码器性能的 *PSNR* 和码率没有影响,使处理速度又提高了一倍。经过算法优化和结构优化,本系统可以达到实时编码的要求。

编码器算法和结构优化前后对 foreman 序列处理过程中各个模块所占用时间的比例如表 5 所示。算法优化后的运动估计模块计算量在整个编码器中的比重有明显下降,结构优化使所有模块的耗时都有所减少。

表 4 编码器优化前后性能比较

序列	优化前			算法优化后			结构优化后
	PSNR/dB	码率/(kbit/s)	帧率/(f/s)	PSNR/dB	码率/(Kbit/s)	帧率/(f/s)	帧率/(f/s)
silent	36.05	543.08	6.32	36.02	569.07	12.57	27.29
news	38.30	456.83	6.80	38.19	477.33	14.25	31.07
mother-daughter	39.11	281.24	6.72	39.02	291.42	14.64	31.19
mobile	34.95	2 569.40	5.57	34.90	2 692.90	12.08	25.60
foreman	36.70	905.37	5.18	36.59	934.12	11.71	24.20
coast-guard	35.81	1 482.10	5.29	35.77	1 545.80	11.31	23.89
football	36.80	2 059.10	4.39	36.70	2 111.80	9.74	20.15

表 5 优化前后各模块耗时及所占比例

模块	优化前		算法优化后		结构优化后	
	耗时/s	比例/%	耗时/s	比例/%	耗时/s	比例/%
1/2 像素插值	3.071	5.14	2.643	10.31	0.121	0.97
运动估计	44.227	74.06	12.818	50.02	2.767	22.31
运动补偿	6.809	11.40	5.879	22.94	3.509	28.29
DCT和 IDCT	2.116	3.54	2.119	8.27	1.875	15.12
量化和反量化	2.733	4.58	2.534	9.89	0.428	3.45
VLC	1.535	2.57	1.484	5.79	1.329	10.71

参考文献

[1] GB/T 200090.2--2006, Information technology-Advanced coding of audio and video-Part2: Video[S].2006.

[2] PSSBK G, KORADA R. MPEG-4 video encoder on ADI blackfin DSP for digital imaging applications[EB/OL]. [2008-04-20]. http://www.geocities.com/ramkishor/.

[3] 任丽香, 马淑芬, 李方慧. TMS320C6000 系列 DSPs 的原理与应用[M]. 北京:电子工业出版社,2000.

[4] 袁义雷, 陈健. 基于 DSP 的 AVS 实时压缩系统设计与实现[J].电视技术, 2005(2):50-52.

[5] 曾庆如, 毕笃彦, 王洪迅. TMS320C64x EDMA 的图像数据传输优化[J]. 电视技术, 2005(S1):66-68.

作者简介:

张 妍(1984-),硕士生,主研数字电视与数字视频;

薛永林 (1965-),副教授,从事数字电视与数字视频的研究;

赵 康 (1983-),硕士生,主研数字电视与数字视频。

责任编辑:任健男 **收稿日期:2008-10-15**

中央电视台与英特尔宣布建立联合实验室

近日,中央电视台和英特尔公司在北京宣布建立联合实验室,宣布双方将共同就多项技术进行联合研发。此次联合实验室的建立将有助于双方进一步加强多媒体处理以及高清视频编辑的能力,并在网络传输、并行计算和数据中心方面进行深入合作。同时,联合实验室作为开放平台将为广电行业的核心应用开发提供技术支持。目前,中央电视台正在进行新台址工艺系统的建设,新台址将建设一个全新的高清网络制播环境来提高高清电视节目制作质量和制作效率。英特尔公司将为中央电视台提供包括基于全新酷睿微架构在内的多项创新技术,并将帮助中央电视台全面提升现有系统的视频处理能力及速度,以满足中央电视台对于高清电视节目制作的要求。

除高清数字电视内容制作外,双方成立的联合实验室将在网络传输、并行计算和数据中心方面进行深入探讨和研究。英特尔将把全球领先的技术和经验,诸如在虚拟化技术的使用,绿色数据中心的建立等方面的技术和经验带到此次与中央电视台的合作中来。这一系列的合作和尝试将进一步提升中央电视台目前的 IT 基础设施建设,从而为中央电视台的技术发展提供动力。