

基于软硬件相结合的多媒体视频终端

毛 讯 王 维 东 姚 庆 栋 虞 露

(浙江大学信息与智能系统研究所, 杭州 310027)

摘 要 随着计算机和网络应用的普及,基于 PC 的多媒体视频终端将具有非常广阔的发展前景.但是由于计算机运算速度的限制,目前还不能完全用软件来实现视频终端的编解码.鉴于这种状况,在对视频编解码算法特点做了详细分析的基础上,提出了一种软硬件相结合的视频终端设计思想.根据该思想设计的视频终端带有一个视频微处理器,它用来处理低级的编解码运算,如进行 DCT/IDCT、ME/MC 以及量化/反量化等,而 PC 机则进行变长编码/变长解码、编解码流的组成与分解、复接/解复接等中高级任务的处理.这样,不仅充分发挥了软硬件的各自优势,使总体性能有了很大提高,而且提供了更多的功能.这种设计思想还为实现 MPEG2 编码器提供了一种更经济的途径.

关键词 视频编码 多媒体视频终端

中图法分类号: TN919 文献标识码: A 文章编号: 1006-8961(2001)06-0599-05

Software/Hardware Based Multimedia Video Terminal Embedded in PC

MAO Xun, WANG Wei-dong, YAO Qing-dong, YU Lu

(Institute of Information and Intelligent System Zhejiang University, HangZhou 310027)

Abstract With the popularity of the application of computer and network technology, there is a bright prospect for the multimedia video terminal embedded in PC. But because of the speed limit of current computer, this terminal could not be realized with software fully. According to this condition, we analyses the computation of video codec in detail, and present a design idea for multimedia terminal based on the cooperation of software and hardware. With this idea, a multimedia video terminal is designed with a video microprocessor which is used for low-level codec computation, such as DCT/IDCT, ME/MC, and Quantization/Inverse Quantization, and PC is used for processing medium and high level task such as VLC/VLD, the construction and deconstruction of coded stream. So it takes each advantage of them fully, and not only enhances the global performance greatly, but also provides more functions. This also indicates a more economic way for MPEG2 encoder realization too.

Keywords Video encoding, Multimedia video terminal

0 引 言

目前 PC 机的计算能力达到 100M 次/s 以上,且软件方式的 MPEG-1 解码也已可实现.而对于质量高的双工视频多媒体会议应用来说,因为编码的复杂度是解码的 4 倍左右,故目前依靠 PC 机用软件来解决还做不到.本文提出一种嵌入于 PC 机中

的辅助计算模块,它是以视频编解码算法为基础的软硬件相结合的模块.通过将视频编解码算法在软件和硬件之间作合理分配,则可达到最高效率.

目前的会议电视产品主要有以色列 VCON 公司的基于 H. 320/ISDN 标准的 C-75 终端和 H. 323 标准的 E-25Pri 终端^[1,2],以及 Microsoft 的 Net-Meeting 等.

表 1 视频编解码算法分类表

算法种类	算法特点
DCT/IDCT, ME, 量化/反量化	运算量大, 占 70% 以上, 算法规则, 可采用并行处理, 运算密集型
变长编解码	运算量中等, 约占 20%, 要存储大码表
H. 263 码流的组成与分解	运算量小, 算法很不规则, 操作密集型

1 视频编解码算法和软硬件相结合的设计思想

1995 年, ITU 通过了关于实现国际连接的会议电视编解码器的 H. 263 标准^[3]. 为了使该建议能在采用不同电视制式的地区之间使用, H. 263 的输入图象格式规定为 CIF(通用中间格式)或 QCIF(四分之一中间格式). CIF 的参数为: 每帧 288 行, 每行 352 个像素, 帧频为 29.97 帧/s. 对于活动图象的信源编码采用运动补偿的帧间预测与 DCT 相结合的混合编码方法. H. 263 建议将视频信号分为图象层(picture layer), 块组层(group of block layer), 宏块层(macro block layer)和块层(block layer), 并且规定一个图象层对应于一帧图象; 一个图象(352×288)分成为 18 个块组, 每个块组又分为 22 个宏块, 每个宏块由 16×16pixel 的 Y 信号和与空间相对应的 8×8pixel 的 Cb, Cr 信号组成.

1.1 H. 263 视频编码器的原理

H. 263 标准的编码是以宏块(MB)为单位进行的, 它的编码方式分为帧内编码和帧间编码, 其在作帧内编码时, 编码器首先对图象的每一个块作离散余弦变换(DCT)^[4]和量化, 然后编码器将量化后的数据一方面输出到视频复合编码器, 另一方面进行反量化和离散余弦反变换(IDCT), 得到本地解码结果, 并将该结果存入帧存中, 作为下一帧图象帧间编码的参考; 而在进行帧间编码时, 视频信号一方面被送到帧存里, 用来进行运动搜索以获得匹配块, 并输出运动矢量, 另一方面计算与前一帧匹配块的帧差, 以减少帧间图象的相关性. 编码器在对帧差信号做完 DCT 变换和量化后, 再将信号输出到视频复合编码器中, 同时又将本地解码的结果存入帧存. 视频编码器输出的信号经过视频复合编码器和码流综合器的处理后就形成标准的 H. 263 码流.

由于计算机速度的限制, 采用全软件的方法不能实时地进行 H. 263 编解码, 所以, 需增加一个视频信号处理器(VSP)来协助 PC 机做编解码工作.

1.2 软硬件相结合的设计思想

为了将编解码任务在 PC 机和 VSP 之间进行合理的分配, 首先需要分析 H. 263 编解码算法的特点. 如今视频编解码器的算法主要分为 3 类(表 1).

其中, DCT/IDCT、ME、量化/反量化等运算较规则, 但计算量大, 因有不少重复运算, 并伴随有许

多次固定的循环, 如果采用 VSP, 则可以大大提高运算效率, 还能减轻 PC 机的负担; H. 263 码流的组成与分解等运算算法非常不规则, 常常有很多分支指令, 若将它排成流水线, 则会产生很多的流水线停顿(stall), 但因其运算量不大, 所以更适于用软件来完成; 而变长编解码运算在采用不同的算法时, 对系统的要求也不同, 因此要做进一步分析. 变长编解码通常有两种算法: 逐位计算法和部分查表法. 前者的特点是运算量大, 速度慢, 且算法不规则, 排成流水线后效率不高; 后者的特点是运算量小, 速度快, 但要存储几个变字长编解码码表, 其中变长编码占 2 104bytes, 变长解码占 3 778bytes, 这对于 VSP 来说是很大的数目, 而对采用虚拟内存的 PC 机来说却是微不足道的. 有一点是肯定的, 即不论采用哪一种算法, 用软件实现都比用硬件实现更合理.

由此可见, 从算法的特点来看, 单纯采用硬件或软件做编解码都不是效率最高和最合理的办法. 本文将这种根据算法的特点, 在软硬件之间进行合理的任务分配的方法称为软硬件相结合的设计方法.

但要实现这种设计方法, PC 机必须要满足以下 2 个条件:

(1) 能执行多任务;

(2) 能提供足够长的空闲时间来进行被安排的编解码运算;

由于当前的个人计算机完全能达到这两个要求, 因此本文设计了相关的视频编解码器.

2 视频编解码器设计与实现

2.1 Axpe1280 介绍

AxPe1280 微处理器是由德国汉诺威大学开发的视频信号微处理器, 该处理器采用了一种由一个 RISC^[7~9]和一个 SIMD 阵列组成的异构 MIMD 结构^[5,6](见图 1 所示).

RISC 微处理器采用 4 级流水线, 即取指令、译码/读操作数、执行和回写等 4 级处理. 其单条指令

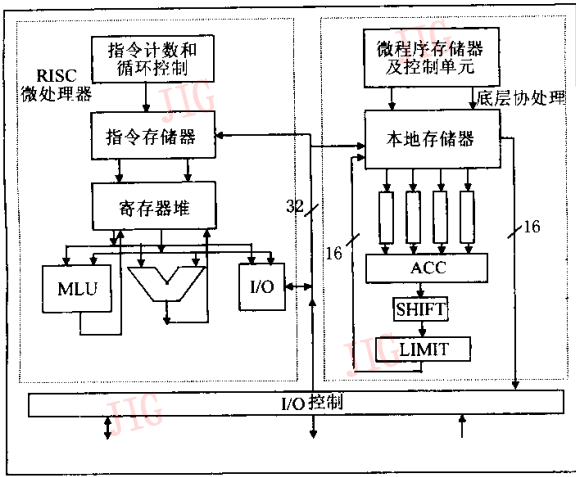


图 1 AxPe1280 的内部结构

的工作过程是:在第 1 级处理时把指令从 2 048×32 bits 指令存储器中取出;在第 2 级处理时作译码并从 256×16bits 寄存器堆中读取操作数;在第 3 级处理时即执行第 1 级处理取出的指令;在第 4 级处理时,将执行结果写回到寄存器堆中.另外,一个专用的 I/O 用来作为寄存器堆和本地存储器的接口.

其底层协处理器由一个 4 096 bytes 的本地存储器、一个微程序控制器和一个包括 4 条数据通路的算术处理单元组成.其算术处理单元如图 2 所示,它是专门为 DCT/IDCT,ME/MC 等底层视频编解码运算设计的.

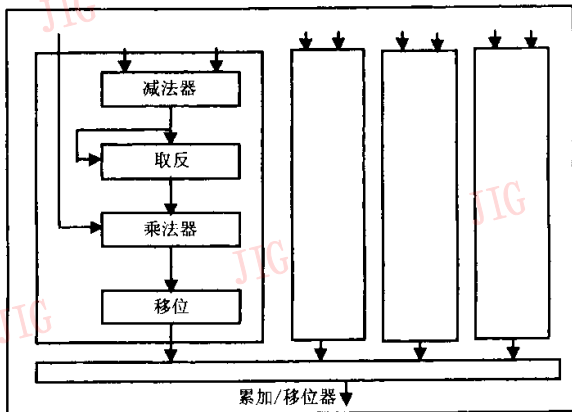


图 2 算术处理单元的数据通路结构

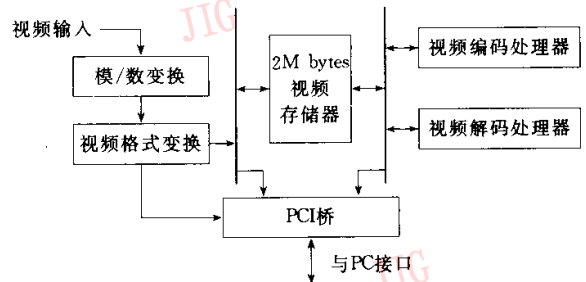
它同时可以进行 4 路运算,其中每一路的典型运算为

$$\frac{1}{2^s} \sum \pm(A-B) \times Coeff \times 2^s \quad s \in (0, 8) \quad (1)$$

式中, A、B 和 Coeff 都是输入参数,其中, Coeff 通常是指 DCT 系数.

2.2 视频编解码器设计

该协处理器作 8bit 运算时,最高运算能力为 1 056MOPS(每秒 1 056M 次操作). AxPe1280 的最高时钟频率可达 66MHz. 视频编解码器的底层编解码部分示于图 3:



从摄像机采集到的模拟全电视信号,经过模数变换和视频格式变换后,以 YUV 4 : 2 : 0 的 CIF 或 QCIF 信号格式存入视频存储器.在做视频编码时,首先用视频编解码处理器对视频数据进行运动估值、DCT 变换、量化等处理,然后将结果按固定格式输出到 PC 机,再由 PC 软件来做变长编码和码流形成工作;解码过程则与编码过程相反,即视频编解码处理器对 PC 机送来的数据进行运动补偿、IDCT 变换和反量化等处理,最后形成 CIF 或 QCIF 格式的图象数据.

视频编解码处理器与 PC 机之间通信的数据以宏块为单位,通信数据包括宏块头信息和 DCT 系数的游程码(run-level 对)等.

视频复合编解码部分的工作在 PC 机上做,当视频编码器每编好一个宏块的数据后就要通过 PCI 桥向 PC 机发送,而 PC 机收到数据后就进行视频复合编码.其复合编码程序是先产生图象层的起始码(PSC),再根据接收到的数据结构来决定图象层、块组层和宏块层的相应头信息.这种头信息分为定长码和变长码两种,其中,由于定长码的码长固定,因此只要根据从 AxPe 获得的头信息数据,在相应的比特位填入适当的码字即可;而对于变长码,为了加快编码速度和减少程序代码,应将不同码表转换成不同的二维数组,而每个数组元素由 2 个分量组成,其中一个代表码值,一个代表码长.编码时,用数组的下标来查得相应的元素,同时将码值写入码流中,并对码流作移位,而位移值即为该码的码长.块层由 DCT 系数组成,DCT 系数采用变长编码,将 DCT 系数按上面的方法制成码表数组,进行查表即可.在 DCT 系数中,对出现概率很小的码字采用换码编码.当出现码表中不存在的码字时,则需要先写入一

个换码:0000011,然后按定长码写入 run 和 level.

在解码时,码流先由 PC 机程序做复合解码,这种复合解码包括对定长码的解码和变长码的解码,其中,定长码的解码比较容易,而变长码的解码有二种方法,一种是逐位比较法,即每次读取一个比特位,然后与码表中的码字作比较来解出码值;另一种是采用分段查表法,由于其解码速度远远高于逐位比较法,所以本文采用分段查表法来解变长码,若根据哈夫曼编码原理,将码表中所有码长小于或等于一个特定长度 L 的码字扩展成长度为 L 的码,则一个码字将与多个扩展码相对应,并且与不同码字相对应的扩展码字集合之间没有交集.根据这个原理,本文将码长分别属于 $L < 7, 7 < L < 10, 10 < L < 12$ 的码字扩展成不同的二维码表数组.这种数组中的每个元素由码值和码长两部分组成,该数组的下标是扩展码的码值.解码时由于是用扩展码来查出对应码字的码长和码值,从而可解出变长码字. DCT 系数的码值由 run 和 level 两部分组成,分别占 8bit 和 4bit.当从码表中解出的是转义码时,需要再读取 6bit 的 run 和 8bit 的 level.

3 码流的发送与接收

软硬件相结合的编解码模式,为用软件来实现编解码流传输提供了无缝接口.充分利用软件的灵活性,即能很容易地使以下两种终端间的通信方式共存于同一个应用程序中.

3.1 符合 H. 320 建议的视频通信

H. 320 建议包括一系列建议,其中,H. 221 建议^[10]定义了 64~1 920kbit/s 信道复用的帧结构^[11],它规定一帧由 80 个连续的 PCM30 帧组成,即为 80×32 时隙,所以其帧率为 100 帧/s,速率为 2Mbit/s.每一帧按数据性质可分为以下几部分:第 0 时隙 (TS_0) 和 TS_{16} 用来保留 G. 704 标准的规定; $TS_2 \sim TS_{15}$ 和 $TS_{17} \sim TS_{31}$ 用来传送视频编解码流; TS_1 用来传送音频编解码流、同步码、信令和数(如文字)等.

大家知道,终端在将数据发送出去之前要将视频编解码流、音频编解码流以及其他相关数据进行复接,以组成帧结构;同样,接收端也要先解复接,然后才能解码,其复接和解复接的方法主要有硬件法和软件法.由于 H. 263 码流最终由软件形成,所以本文采用软件的方法来复接和解复接.对于复接,软件要做的工作是将视频/音频的编解码流、同步码、

信令和数(如文字)等填入帧存(相当于一帧大小的连续内存)的相应位置,并且每隔 10ms 将其更新一次,而传输则采用 DMA 方式将这段内存循环读出,读出数据是严格按 2Mbit/s 进行的;对于解复接,要做的工作正好相反.

在复接和解复接中,要处理好读写的时间,比如,在复接时,如果数据更新太慢则会造成数据重复发送,相反,如果太快,则又会丢失数据.为了解决这个问题,本文定义了 2 个帧存,并使这两个帧存轮流处于更新数据或读取数据的状态.同时,在 DMA 处理上,本文又定义了一个事件来标志帧存的状态变化.这样,当事件发生时,一个自定义的消息就会被发送出去,用来触发相应的处理机制.这样即可实现数据的异步发送,从而大大提高了程序的运行效率.解复接也采用类似的方法.

3.2 基于 H. 323 的视频通信

1996 年国际电信联盟 (ITU-T) 公布了 H. 323^[12] 建议,该建议定义了 LAN 上进行多媒体通信的设备、规程和协议.目前,会议电视设备日趋微机化、多功能化,并已使得数据会议与视讯会议技术相融合.由于该技术是基于计算机的,所以它的成本可以大大降低,并与其他媒体很好地在计算机中融合.

由于本文的编解码流是由软件最终形成的,所以要建立这种通信机制非常容易,而本文就是调用了关于 Windows Socket 的 API 函数来实现,该工作首先要在收发双方创建基于数据流的 SOCKET,即一方作为服务器 (Server),另一方作为客户机 (Client),并且分别申请接收和发送缓冲区;接着,Client 要调用 connect() 来建立连接,一旦 server 响应 connect(),则双方就可以开始通信.在通信时,发送端将产生的 H. 263 编解码流数据存入发送缓冲区,一旦 FD_WRITE 网络事件发生,则调用 send() 将编解码流从发送缓冲区发出;在接收端,一旦 FD_READ 网络事件发生,则调用 recv() 将数据存入接收缓冲区,用来做解码.本文还采用了异步通信方式来提高效率,并用多线程技术,以保证在作图象编解码和数据传输的同时,能够响应用户输入.

显然,如果不用软硬件相结合的方法,该编解码器就不可能实现基于 TCP/IP 的视频通信.

4 结 论

本视频终端与 Microsoft 的 NetMeeting 的性

能比较结果如表 2。

表 2 本视频终端与 Microsoft NetMeeting 的性能比较

	占用计算机系统运算资源	帧率(帧/s)	运行大程序时的帧率(帧/s)
本视频终端	10%~23%	30	25~27
NetMeeting	>60%	12~20	5~10

通过比较可看出, NetMeeting 将占用 60% 的系统运算资源, 而本视频终端只占 10%~23%。两者的平均帧率相差也比较大, 当计算机在运行较大的应用程序比如 Visual C++, Office, Matlab 时, NetMeeting 的帧率会降到 5 帧/s, 而本视频终端的帧率能保持在 25 帧/s 以上, 并具有较好的图象质量。

通过前面的介绍和实践可看出, 基于软硬件相结合设计思想的多媒体视频终端至少具有以下优点:

(1) 编解码器的可编程性, 大大提高了它的灵活度和可操作性, 同时也便于升级。如将原有的 H. 263 编解码器升级为支持高级编码模式时, 只要将 PC 和 AxPe1280 程序作部分修改即可, 这样不仅节省了大量的开发时间, 而且不必再花费其他成本;

(2) 复接和解复接操作也可采用软件来实现, 这样既可降低成本, 同时也为视频复合编解码与复接/解复接提供了无缝隙接口, 从而可减少调试的难度, 提高效率, 缩短开发周期;

(3) 有很强的开放性。本视频终端不仅能实现基于 H. 221 传输标准的会议电视, 而且对 TCP/IP 也提供了支持, 还可以实现基于 Internet 的视频多媒体通信;

(4) 为 MPEG2 编解码器的设计提供了一种新的方案。采用软硬件相结合的设计方法可使编解码协处理器与 PC 机来协同进行 MPEG2 编码, 而由于编解码协处理器只做编解码运算的一部分工作, 所以对其性能要求不高, 价格也会比较低廉。

目前, 本视频终端已达到 H. 263 标准的要求, 并且在我校为德国汉诺威大学开发的 H. 263 高级编码模式项目中进行了成功的演示。

致谢 感谢汉诺威大学的 Mussmann 教授和 Jörg Hilgenstock 博士为我们提供了 AxPe1280 芯片模板及开发工具。同时, 许多问题在和他们的共同讨论中得到解决。

参 考 文 献

- 1 朱秀昌、刘峰. 会议电视系统及应用技术. 北京: 人民邮电出版社, 1999.
- 2 黄东霖. 多媒体业务的主流 IP 网上的多媒体通信业务. 世界网

络与多媒体, 1998, 6(1), 8~13.

- 3 ITU-T Recommendation H. 263. Video Coding for Low Bitrate Communication. Geneva, 1996.
- 4 姚庆栋, 毕厚杰, 王兆华等. 图象编码基础. 杭州: 浙江大学出版社, 1993.
- 5 PHILIPS Corp. AxPe1280V Programmable Video Signal Processor Data Sheet. 1998;144~58.
- 6 PHILIPS Corp. AxPe1280V Programmable Video Signal Processor Data Sheet. 1998;2;5~6.
- 7 Hermann K, Seifert M, Gaedke K *et al.* A RISC core for a monolithic video signal processing, New York: IEEE, Oct. 1994;368~377.
- 8 Araki T, Toyokura M, Akiyama T *et al.* Video DSP architecture for MPEG2 codec. IEEE International conference on acoustics, speech and signal processing. ADELAIDE, SOUTH AUSTRALIA. April 19-22, 1994;417~420.
- 9 Pirsh P, Stolberg H-J. VLSI implementations of image and video multimedia processing systems. IEEE Transaction on circuits and systems for Video Technology. 1998, 8(7);878~891.
- 10 ITU-T Recommendation H. 221. frame structure for 64 to 1920 kbit/s channel in audiovisual teleservices, Geneva, 1995.
- 11 郑世宝. 会议电视技术综述. 通信技术与发展, 1998, (4):1~10.
- 12 ITU-T Recommendation H. 323. Visual telephone systems and terminal equipment for local area networks which provide a non-guaranteed quality of services. Geneva, 1995.

毛 讯 1970 年生, 现为浙江大学信息与智能系统研究所博士生. 主要研究兴趣包括图象编码、数字图象通信和视频微处理器设计。

王维东 1967 年生, 现为浙江大学信息与智能系统研究所博士生. 主要研究兴趣包括图象编码、数字图象通信和视频微处理器设计。

姚庆栋 1932 年生, 浙江大学信息与智能系统研究所教授, 博士生导师. 主要研究兴趣包括 VLSI 设计、图象编码。

虞 露 1969 年生, 浙江大学信息与智能系统研究所副教授. 主要研究兴趣包括图象编码、数字图象通信。