

基于局域网的视频解码算法的研究与实现

彭波 胡治国 吕小晴

(中国农业大学 信息与电气工程学院,北京 100083)

摘要 网络视频解码的基础是码流的管理和视频流的解析。所设计的基于局域网的视频解码算法采用了码流的三级缓冲管理机制:由接收缓冲区负责存储从服务器端接收到的码流,工作缓冲区直接向视频解码提供所需数据,中间缓冲区负责以上2个缓冲区的中转。码流的解析首先从复合流中提取视频流,然后从视频流中读取一定位数的数据进行解码。为了保证视频解码过程的随机开始,在解码开始前,要求服务器先发送有关信息以便得到解码所需的必要参数,再接收码流并从找到的第1帧I图像开始解码。为了正确解出紧跟在第1帧I图像后面的B图像,系统构造了前向预测图像数据。抖动处理提高了图像显示的质量和速度。

关键词 MPEG-1; 视频; 码流; 解码; 算法

中图分类号 TP 301.6; TP 311

文章编号 1007-4333(2004)02-0073-05

文献标识码 A

Study and implementation of video decoder algorithm based on LAN

Peng Bo, Hu Zhiguo, Lü Xiaoqing

(College of Information and Electricity Engineering, China Agricultural University, Beijing 100083, China)

Abstract The management of the bitstream and the parsing of video data are the basis of the video decoder. Video decoder algorithm based on LAN adopts three kinds of buffers were offered to store the bitstream. The receiving buffer stores the bitstream come from the server. The working buffer directly gives needed bits to the decoder. The middle buffer associates the above-mentioned two buffers. The bitstream parsing consists of two steps. Firstly the video data must be extracted from the compound bitstream, then required bits are provided for the decoder. Since the decoding was asked to begin at any time, the client requests the server to transfer the information to fill the indispensable parameters before receiving the bitstream, furthermore the decoder ignores all the bitstream until the first I frame comes. In order to decode the B frame following the first I frame, the algorithm creates a previous prediction picture. The Bayer ordered dithering was used to show the sequential pictures nicely. The algorithm have been used to video monitor system of bank, and have good stability. Offers a satisfied resolution for the similar applications.

Key words MPEG-1; video; bitstream; decoder; algorithm

基于局域网的新一代远程网络视频监控系统由前端设备、服务器和客户端3部分组成,其中,客户端的主要功能是实现远程监控。要实现远程监控,就必须完成实时视频解码。从降低成本的角度考虑,软件实现是一种很好的选择,而现代微型计算机的强大处理能力可以满足软件视频解码的需求。

目前国内外开发的影音播放器都可以完成基于文件的解码,如美国 Microsoft 公司的 Windows Media Player,北京豪杰公司的“超级解霸”等,但它们只是从计算机现存的文件中读取数据进行解码,不适

合实时网络视频解码的要求。利用美国 Microsoft 公司提供的 DirectShow 技术,可以实现实时 MPEG-1 数据解码的功能,但是当切换通道时,会产生错误从而使系统崩溃。另外,由于 DirectShow 的内部核心技术不公开,因而无法对其进行改进,所以该技术也无法运用于实际的网络视频解码系统。国内北京黄金眼公司实现的远程监控系统、北京联成阜康科技公司实现的 nDVR 网络视频监控系统等,都具有自主知识产权的产品,其实现技术不公开。

有鉴于此,开发一个网络视频解码算法是必要

收稿日期: 2003-07-01

作者简介: 彭波,副教授,主要从事多媒体与网络技术的应用等研究。

且有价值的。笔者设计的基于局域网的视频解码算法具有如下功能:客户端从服务器端接收码流,正确解码图像并实时显示;图像质量清晰,显示过程流畅;解码能够随时开始、随时结束,中途不能出错退出;实现1路和4路通道码流的同时解码和显示。

1 网络视频解码系统的总体结构

网络视频监控分为服务器端和客户端2个部分。服务器端主要负责数据采集和编码,以及侦听客户端连接要求后的数据发送。客户端主要完成从服务器端接收实时码流数据送给解码系统,经过数据缓存,根据需要进行视频数据的解析,解码出的数据经抖动处理进行图像显示。视频解码算法采用静态库(library)的结构形式,开发环境为 Visual C++ 6.0^[1]。图1为视频解码系统工作示意图。

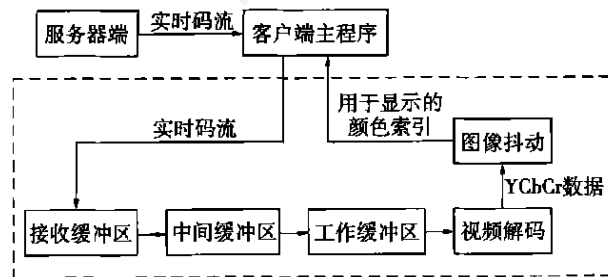


图1 视频解码系统工作示意图

Fig. 1 The operation of the video decoder system

视频解码系统分为4部分:

1) 码流的缓冲管理。本系统要实现的解码不同于从文件中读取数据的解码,其码流数据从网络上实时接收,接收、解码及显示同时进行。本系统采用码流的三级缓冲管理机制,即:用于从服务器端接收码流数据的接收缓冲区,直接服务于解码过程的工作缓冲区和保证平稳解码而设计的中间缓冲区。良好的码流缓冲管理为网络视频解码系统的成功奠定了基础。

2) 视频流的解析。包括2个方面,即从复合流中得到视频流,从视频流中取得数据为解码过程服务。视频解码过程按照MPEG视频的语法进行,根据语法决定码流的读取。处理完一个语法单元,再处理下一个,在需要相关信息时,从原始码流中解析得到。

3) 视频的解码。负责将MPEG-1视频数据解码为连续的YCbCr图像数据^[2]。视频解码以图像为中心展开,在解码完图像中的所有宏块之后,就可以重构出图像数据。宏块的解码,需要经过变长码

(variable-length code)解码和行程码(run-length code)解码得到DC系数和AC系数;经过反扫描,恢复成8×8矩阵中按顺序排列的变换系数;经过反量化处理,再进行反余弦变换,如果是预测图像,还需要经过运动补偿处理。

4) 图像的显示。将YCbCr格式的图像数据转换为可由计算机直接显示的图像数据。本系统采用图像抖动处理技术完成转换,很好地解决了图像显示质量与速度之间的矛盾。

综上所述,客户端解码的工作过程为:由主程序负责从服务器端接收码流数据传给解码系统,解码系统经过数据缓存,为解码过程提供数据服务,解码出的图像数据经过抖动变换后再传回到主程序端显示。

2 关键技术及其实现

通常的MPEG解码器工作时,是从本机存储的文件中取得数据进行解码的,由于其数据读取是按顺序进行的,因此不会出现数据被跳过或数据出错的情形。而本系统要实现的解码,其码流数据从网络上接收,要求实时解码和显示,其情况完全不同于从文件读取数据的解码过程。

2.1 码流管理机制

由于网络环境的不稳定性,要实现顺利解码,提供可接受的视频质量,必须有良好的码流缓冲机制以提供类似文件的稳定的数据服务。本算法设计了三级码流缓冲管理机制。

1) 接收缓冲区。一旦客户端连接上服务器端,就会不断地接收到实时码流,这些数据不可能被马上处理完,因此必须有一个缓存区,这就是码流接收缓冲区。由于单位时间内接收到的数据量比较大,因此在内存的堆(heap)存储区中开辟了一个空间作为接收缓冲区。该缓冲区既要接收码流,又要为其他解码缓冲区提供码流,因此本系统中设计了一种循环队列数据结构^[3]管理数据的存取。

2) 工作缓冲区。码流是按照MPEG-1语法规则组织的,解码过程要求的信息都必须从码流中得到,因此解码过程首先要解析码流,以便得到解码控制信息和获取运算数据,这个经常性的处理决定着解码的速度。这就要求有一个高效的码流缓冲区来提供数据,因此在存取速度较快的栈(stack)中开辟了一个空间作为工作缓冲区。

3) 中间缓冲区。由于网络的不稳定性,接收码

流数据或快或慢,为了保证解码的稳定性和解码图像的连续性,在接收缓冲区与工作缓冲区之间构造了一个过渡性的缓冲区——中间缓冲区,可以屏蔽接收端,为工作缓冲区提供稳定的数据服务。

4) 3 种缓冲区的配合工作机制。当解码过程需要新的码流数据时,从工作缓冲区中读取。如果工作缓冲区中没有数据,则先从中间缓冲区中读取数据,再进行解码。如果中间缓冲区中也没有未使用过的数据,则先从接收缓冲区中读取数据,再供给工作缓冲区。如果接收缓冲区为空,则分为 2 种情况处理:一是主程序要求结束解码过程,此时就在工作缓冲区中填充序列结束代码,解码过程读取到该代码时自动结束;二是主程序没有要求解码过程结束,这说明网络接收数据太慢,此时暂停解码过程,等数据到来后再继续解码。

5) 3 种缓冲区的大小。工作缓冲区直接为视频解码过程提供数据,由于 MPEG1 视频数据是以包(packet)为单位组织的,而通常一个视频包的大小为 2 048 Byte,因此设计工作缓冲区大小为 2 048 Byte。

中间缓冲区是联系接收缓冲区和工作缓冲区的纽带。为方便工作缓冲区读取数据,设计中间缓冲区每个缓冲单元的大小与工作缓冲区相同。为了使每次充满数据的中间缓冲区能够为工作缓冲区提供若干次数据,以便可以稳定地解码出几帧完整图像,设计中间缓冲区的大小为 $16 \times 2\,048$ Byte,可供工作缓冲区读取数据 16 次。

受网络状况的影响,每次接收到的数据大小可能有所不同,为避免频繁存储不同大小的数据,可以先凑足一定的数据量再进行存储,这样就构成了接收缓冲区的一个存储单元。为便于中间缓冲区读取数据,每个单元大小设计为 $32\,768(16 \times 2\,048)$ Byte。存储单元不能只有 1 个,否则数据容易覆盖,也不能太多,否则太占用内存。这里设计其大小为 $16 \times 32\,768$ Byte,可供中间缓冲区读取 16 次。

所设计的缓冲区大小,是在考虑完成各个缓冲区功能的基础上,结合解码的具体情况确定的经验值。实际测试表明,设计是合理的。

2.2 解码系统中各个类的设计

根据解码系统的工作流程,采用面向对象的设计思想,在解码静态库中设计了 5 个 C++ 类,它们之间的相互调用关系见图 2。

1) CDecodeControl 类:负责与客户端主程序的

交互。所有主程序与静态库的接口都通过该类提供,从而屏蔽了内部实现的细节。主要提供 2 类接口:控制解码和传递码流数据。控制解码包括启动解码和停止解码。传递码流数据是将主程序提供的实时码流传递给接收缓冲区缓存。

2) CReceiveBuffer 类:接收缓冲区的管理,即缓存接收到的数据,并给中间缓冲区提供数据。

3) CVideoDecode 类:负责 MPEG 视频的解码和图像数据的抖动处理,重构出图像。

4) CBitStream 类:实现和管理工作缓冲区,即工作缓冲区从中间缓冲区中读取数据,供解码过程使用。

5) CMidBuffer 类:管理中间缓冲区,即当工作缓冲区的数据不足,并请求数据时,给工作缓冲区提供数据;当自己没有可用数据时,从接收缓冲区中获取数据。

2.3 视频解码的实现

MPEG1 中存在 3 种类型的图像:I 图像、P 图像和 B 图像。其中 I 图像仅根据自身的信息就可以正确解码,称为帧内图像;P 图像和 B 图像称为预测图像,必须借助于前面或后面图像的信息才能进行正确解码。

1) 解码的随机开始。由于客户端与服务器的连接是随机的,因此它接收到的第 1 帧图像不一定是 I 图像,这样将导致无法正确解码。本系统中采用以下设计方案来解决该问题:检测到第 1 帧 I 图像后,才开始解码图像。由于每隔一定帧数(通常在 10 帧左右)视频流中就会出现 1 个 I 图像,因此这样处理不会跳过太多的数据。

由于解码 1 帧图像除了需要图像数据本身之外,还要求一些其他信息,比如图像的宽度和高度等,而这些信息是在视频序列头(sequence header)中给出的。在视频码流中,所有视频序列头中的信息都是一样的,它们被间隔安排在码流中,以便能够随机访问。由于视频序列头的出现频率是不定的,而等待接收到视频序列头之后才开始解码图像,将有可能损失掉很多数据,故本系统采取了一种方法既避免等待,又使图像解码能够顺利进行,即在初始化解码阶段要求服务器端先发送一个完整的视频序列头过来,然后再发送其他编码数据。

经过以上处理,接收到的第 1 帧 I 图像就可以正确解码了。随后的 P 图像可以根据解码出的 I 图像数据进行解码,而解码出的 I 图像和 P 图像数据

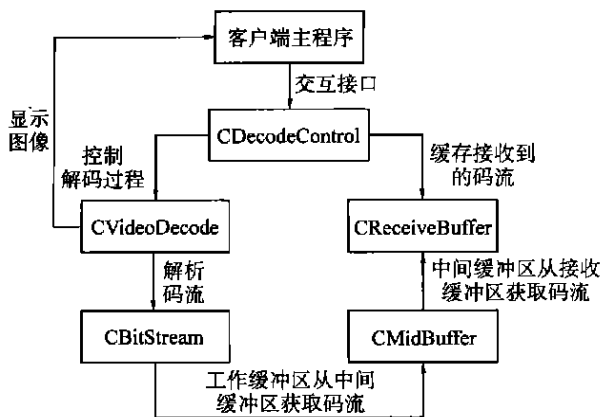


图 2 程序中各类之间的调用关系

Fig. 2 The transfer relation between classes in programming

又为后面的 B 图像提供了必要的解码信息。依此类推,除了紧跟在第 1 帧 I 图像后面的 B 图像无法正确解码外,后面的所有图像都可以顺利解码。

2) 紧跟在第 1 帧 I 图像后的 B 图像的解码。B 图像是双向预测图像,它需要有前向和后向预测图像的解码数据才能正确解码。当编码顺序为“IBB PBB PBBP...”时,紧跟在第 1 帧 I 图像之后的 2 个 B 图像,除了前面解码出的 I 图像数据作为其后向预测图像外,没有前向预测图像,因而无法正确解码。笔者分析认为,此时可以有 2 种处理方法:一种是只利用后向预测图像数据对其进行解码;另一种是“构造”一个前向预测图像。本系统采用了第 2 种方案,即复制已得到的 I 图像数据作为前向预测图像数据。这样处理保证了解码程序的统一性,即 B 图像在所有情况下都是利用了前向和后向预测图像进行解码的。

3) 宏块的解码^[4]。视频解码是连续图像的解码,其中编码宏块的解码又是图像解码的核心。图 3 给出了编码宏块的解码流程。由于宏块是由块组成的,所以图 3 中除了运动补偿以宏块为基本单位进行外,其他处理都是以块为基本单元进行的。

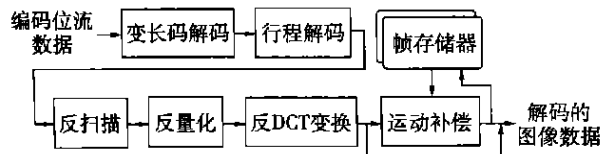


图 3 编辑宏块的解码流程

Fig. 3 The decoder flow chart of the encoded microblock

解码是编码的逆过程,编码位流数据通过变长码解码和行程解码得到块的 DC 系数和 AC 系数;经过反扫描,恢复成 8 × 8 矩阵中按顺序排列的变换

系数;码流中的系数在编码时经过了量化处理,因此还需经过反量化处理,才能进行反 DCT 变换(反余弦变换)。此后分 2 种情况区别处理:

非内部编码宏块的解码。由于非内部编码宏块的编码是本图像与参考预测图像的残差值,因此解码还需利用已经解码出的参考图像,将经过反余弦变换后的数据加上参考图像相应位置的数据才可以得到本宏块的数据,这一过程称为运动补偿。

内部编码宏块的解码。内部编码宏块的编码是宏块本身的数据,因此经过反余弦变换后得到的数据就是本宏块的数据。这种情形不需要进行运动补偿。

2.4 图像的显示

可以在计算机上直接显示的图像数据有 2 种:RGB 格式数据和调色板数据。图像解码得到的是 YCbCr 格式数据,因此必须进行变换后才能显示。

由于 YCbCr 格式的数据变换成 RGB 格式所需的计算量很大,而且一副图像的 RGB 数据本身数据量也很大,显示很慢,无法满足实时要求。因此本系统采用调色板方式显示。如果采用 16 色或 256 色显示,图像会很粗糙,很多细节无法表现出来,因此这种方式不宜在实用系统中采用。

本系统采用图像抖动技术将 YCbCr 数据变换成典型颜色的索引,结合相应的调色板实现了图像的显示。计算量不大,图像的数据量减小了 1/3,而且图像质量也令人满意。其原理是,利用多种可见颜色的组合来模拟一种不可见的颜色,从而达到用较少的颜色来表现较大的颜色空间的目的。算法中用位组合法实现了基于 4 × 4 抖动矩阵的 Bayer 顺序抖动。

静态库中经过抖动处理得到的图像数据,以消息发送的形式传回给客户端主程序显示。

3 结束语

所设计的视频解码算法具有 3 个主要特点:

- 1) 采用了三级码流缓冲机制,为实时解码提供了稳定的数据服务;
- 2) 通过先从服务器端获得解码所需要的必要信息和从找到的第 1 帧 I 图像开始解码的方法,系统可以随时响应开始解码的要求;
- 3) 采用图像抖动处理技术将 YCbCr 数据变换成典型颜色的索引,再结合相应的调色板实现图像的显示,可以很好地解决图像显示质量与速度之间

的矛盾。

本算法已经用于中国农业银行柜员制的网络视频监控系统中,实际运行表明视频解码稳定性良好。在纠错方面,该算法只是保证在发现码流错误时,解码仍然正常进行;但由于对错误没有进行纠正处理,因此图像仍然可能出现局部错误。为进一步完善该算法,在今后的研究中应考虑采取增强纠错功能。

参 考 文 献

[1] Microsoft. MSDN Library Visual C++ 6.0 版. <http://www.microsoft.com/msdn>. 2000

- [2] 马小虎. 多媒体数据压缩标准及实现[M]. 北京:清华大学出版社, 1996. 272~293, 329~414
- [3] 严蔚敏. 数据结构[M]. 北京:清华大学出版社, 1992. 63~65
- [4] 钟玉琢. MPEG2 运动图像压缩编码国际标准及 MPEG 的新进展[M]. 北京:清华大学出版社, 2002. 102~172
- [5] 何志强. 图像抖动技术的原理及实现[J]. 计算机系统应用, 1999(7): 43~46

中国农业大学科技查新工作站简介

科技查新是为科研项目的评审专家提供全面准确的“鉴证性客观依据”,配合专家评议,对科研项目做出客观、公正、准确的评价;并促进各级科研管理部门提高科研立项、成果鉴定与奖励的严肃性、公正性、准确性和权威性。

科技查新的最终结果是针对查新项目向委托人提供一份包括查新过程陈述和新颖性论证的书面报告,即《科技查新报告》。

查新咨询服务对象范围:

- 1) 科研课题立项
- 2) 科研成果鉴定
- 3) 科研成果报奖
- 4) 专利申请
- 5) 技术咨询(包括技术引进、技术转让、新产品开发等)

中国农业大学图书馆自 1990 年开始从事科技查新工作,1995 年被农业部正式批准为首批部级科技查新单位。2003 年 11 月被教育部认定为教育部部级科技查新工作站,成为全国惟一农学类教育部部级科技查新工作站。同时被授予教育部科技查新工作站的还有清华大学、北京大学、复旦大学等全国 29 所知名大学。

10 多年来中国农业大学科技查新工作站共完成国家级、部级等科技查新项目 700 多项,课题专业涉及农业、生物、食品、机械、水利工程、电气等。

目前本馆查新机构具备较强的文献资源保障能力,具有一支经验丰富、专业背景齐全的查新队伍,具有健全的管理制度和规范的工作流程。

本馆查新服务面向校内外广大教师、科研人员及研究生,欢迎大家光临和咨询。本部将以热情、高质量的服务接待每一位用户。

办公地点:中国农业大学图书馆信息咨询部(东馆 301 房间,西馆二楼信息部)

联系人:左文革 电话:62891045 E-mail:zuowg@cau.edu.cn(西馆)

联系人:袁永翠 电话:62336502 E-mail:wyyys@cau.edu.cn(东馆)

(中国农业大学图书馆供稿)